

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-203029

(43)Date of publication of application : 18.07.2003

(51)Int.Cl.

G06F 13/00
G06F 12/00
G06F 15/177
H04L 12/56

(21)Application number : 2002-168620

(71)Applicant : NEC CORP

(22)Date of filing : 10.06.2002

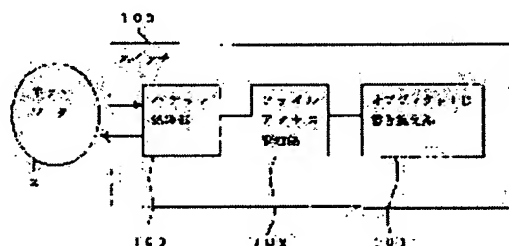
(72)Inventor : YAMAKAWA SATOSHI
ISHIKAWA JUN
TORII TAKASHI

(30)Priority

Priority number : 2001337838 Priority date : 02.11.2001 Priority country : JP

(54) SWITCHING METHOD, DEVICE AND PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a service of single system image to a client by integrally managing two or more servers without tempering with an existing system.**SOLUTION:** An object ID rewrite part 101 rewrites an original object ID, when the original object is contained in a response to be transmitted, to an information-attended object ID by inserting server identification information to the original object ID, and restores the information-attended object ID, when the information-attended object ID is contained in a request to be transferred, to the original object ID. A file access management part 102 has a table for managing the combination of address information for a server and server identification information, determines the address information of the server corresponding to the server identification information contained in the information-attended object ID of the request to be transferred in reference to the table, and determines the transfer designation of the request on the basis of the address information.

LEGAL STATUS

[Date of request for examination] 10.06.2002

[Date of sending the examiner's decision of rejection] 28.06.2006

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-203029

(P2003-203029A)

(43) 公開日 平成15年7月18日 (2003.7.18)

(51) Int.Cl.	識別記号	F I	テ-マ-ト (参考)
G 0 6 F 13/00	5 2 0	G 0 6 F 13/00	5 2 0 C 5 B 0 4 5
12/00	5 1 4	12/00	5 1 4 E 5 B 0 8 2
	5 4 5		5 4 5 A 5 K 0 3 0
15/177	6 7 4	15/177	6 7 4 A
H 0 4 L 12/56	1 0 0	H 0 4 L 12/56	1 0 0 Z
審査請求 有 請求項の数65 O L (全 43 頁)			

(21) 出願番号 特願2002-168620 (P2002-168620)

(22) 出願日 平成14年6月10日 (2002.6.10)

(31) 優先権主張番号 特願2001-337838 (P2001-337838)

(32) 優先日 平成13年11月2日 (2001.11.2)

(33) 優先権主張国 日本 (J P)

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 山川 聡

東京都港区芝五丁目7番1号 日本電気株式会社内

(72) 発明者 石川 潤

東京都港区芝五丁目7番1号 日本電気株式会社内

(74) 代理人 100088328

弁理士 金田 暢之 (外2名)

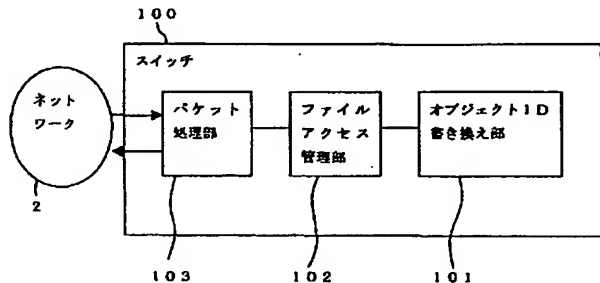
最終頁に続く

(54) 【発明の名称】 スイッチ方法、装置およびプログラム

(57) 【要約】

【課題】 既存システムに手を加えることなく、複数のサーバを統合管理してクライアントにシングルシステムイメージのサービスを提供する。

【解決手段】 オブジェクトID書き換え部101は、転送する応答に、元オブジェクトIDが含まれている場合には、元オブジェクトIDにサーバ識別情報を挿入することによって情報付帯オブジェクトIDに書き換え、転送する要求に、情報付帯オブジェクトIDが含まれている場合には、情報付帯オブジェクトIDを元オブジェクトIDに復元する。ファイルアクセス管理部102は、サーバのアドレス情報とサーバ識別情報との組み合わせを管理するテーブルを有し、転送する要求の情報付帯オブジェクトIDに含まれるサーバ識別情報に対応するサーバのアドレス情報を、そのテーブルを参照して求め、そのアドレス情報に基づいて要求の転送先を決定する。



【特許請求の範囲】

【請求項1】 少なくとも1つのクライアントと少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置されたスイッチ装置における、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ方法において、

転送する応答に、前記オブジェクトを識別するために前記サーバで生成された元オブジェクトIDが含まれている場合には、前記元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報を挿入することによって前記元オブジェクトIDを情報付帯オブジェクトIDに書き換える段階と、

転送する要求に、前記情報付帯オブジェクトIDが含まれている場合には、前記情報付帯オブジェクトIDを前記元オブジェクトIDに復元する段階と、

前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理するテーブルを参照して、転送する要求の情報付帯オブジェクトIDに含まれる前記サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する段階と、

を備えることを特徴とするスイッチ方法。

【請求項2】 前記サーバ識別情報には、オブジェクトが公開される単位である共有オブジェクト単位を識別するための共有オブジェクト単位識別情報が含まれる請求項1記載のスイッチ方法。

【請求項3】 前記要求および前記応答を関連づけ、前記要求および前記応答を前記クライアントおよび前記サーバで区別するために付与されているトランザクションIDとクライアントのアドレス情報との組み合わせを管理する第2のテーブルを参照して、転送する応答に含まれるトランザクションIDに対応するアドレス情報を求め、当該アドレス情報に基づいて当該応答の転送先を決定する段階をさらに備える請求項1または2記載のスイッチ方法。

【請求項4】 前記要求および前記応答を関連づけ、前記クライアントと前記サーバの組み合わせを一意に特定するのに利用可能なOSモデル第4層以下のタグ情報とクライアントのアドレス情報との組み合わせを管理する分類テーブルを参照して、応答を転送すべきクライアントのアドレス情報を求め、当該アドレス情報に基づいて当該応答の転送先を決定する段階をさらに備える請求項1または2記載のスイッチ方法。

【請求項5】 前記要求および前記応答を関連づけ、前記クライアントと前記サーバの組み合わせを一意に特定するのに利用可能なサーバとの通信セッション・ポート

番号とクライアントのアドレス情報との組み合わせを管理する分類テーブルを参照して、応答を転送すべきクライアントのアドレス情報を求め、当該アドレス情報に基づいて当該応答の転送先を決定する段階をさらに備える請求項1または2記載のスイッチ方法。

【請求項6】 前記サーバより公開されたオブジェクトのディレクトリツリーを組み合わせ、1つの統合ディレクトリツリーとして管理する擬似ファイルシステムを用いる、請求項1乃至5のいずれか1項記載のスイッチ方法。

【請求項7】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令情報との組み合わせを管理する第4のテーブルを参照する段階と、転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を、前記各送信先に転送する要求に分割し、前記各送信先から受信した全ての応答を、前記クライアントに転送する応答として合成するファイルシステム補完段階と、をさらに備える請求項6記載のスイッチ方法。

【請求項8】 前記第4のテーブルに登録されるオブジェクトIDは、前記情報付帯オブジェクトIDである請求項7記載のスイッチ方法。

【請求項9】 前記サーバが公開するオブジェクトが移動したときの、移動前における当該オブジェクトの移動元オブジェクトIDと、移動後における当該オブジェクトの移動先オブジェクトIDと、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管理する第5のテーブルを参照する段階と、前記擬似ファイルシステムによって管理されている統合ディレクトリツリーを参照し、所定の基準に基づいて前記オブジェクトを移動させるとともに、当該オブジェクトの移動元オブジェクトIDと、移動先オブジェクトIDと、移動先サーバ識別情報とを前記第5のテーブルにエントリする段階と、

転送する要求に含まれるオブジェクトIDが前記第5のテーブルの移動元オブジェクトIDと一致する場合には、当該オブジェクトIDを、一致した当該移動元オブジェクトIDに対応する移動先オブジェクトIDに変更し、当該移動元オブジェクトIDに対応する移動先サーバ識別情報に基づいて送信先のサーバを決定する段階と、をさらに備える請求項6乃至8のいずれか1項記載のスイッチ方法。

【請求項10】 前記第5のテーブルには、移動元オブジェクトIDとして、前記情報付帯オブジェクトIDがエントリされる請求項9記載のスイッチ方法。

【請求項11】 前記統合ディレクトリツリー管理において特殊操作の必要な命令情報との組み合わせを管理す

10

20

30

40

50

る第4のテーブルを参照する段階と、
転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を他の要求とは異なる通信セッションポート（ポート番号X）を用いて送信し、前記各送信先から受信した応答を、その他の応答から当該ポート番号Xにより識別し、前記クライアントに転送する応答として復路の特殊処理を行う段階と、とをさらに備える請求項6記載のスイッチ方法。

【請求項12】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令情報との組み合わせを管理する第4の管理テーブルを参照し、

転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を、前記各送信先に転送する要求に分割し、当該分割された要求を他の要求とは異なる通信セッションポートに対応するポート番号を用いて送信し、前記各送信先から受信した全ての応答を、その他の応答から当該ポート番号X、Y・・・により識別し、前記クライアントに転送する応答として合成するファイルシステム補完段階と、をさらに備える請求項11記載のスイッチ方法。

【請求項13】 前記統合ディレクトリツリー管理において特殊操作の必要な命令情報との組み合わせを管理する第4のテーブルを参照する段階と、

特殊操作を必要とする要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、OS Iモデル第4層以下で判断できるタグ情報を当該要求に付与してこれを送信し、前記各送信先から受信した応答を、当該タグ情報により識別し、前記クライアントに転送する応答として復路の特殊処理を行うファイルシステム補完段階と、をさらに備える請求項6記載のスイッチ方法。

【請求項14】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令情報との組み合わせを管理する第4のテーブルを参照する段階と、

転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を前記各送信先に転送する要求に分割し、当該分割された要求にOS Iモデル第4層以下で判断できるタグ情報を付与してこれらを送信し、前記各送信先から受信した全ての応答を、当該タグ情報により識別し、前記クライアントに転送する応答として合成するファイルシステム補完段階と、をさらに備える請

求項6記載のスイッチ方法。

【請求項15】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトの属性に、当該オブジェクトがつなぎ目であることを示す属性を記録する段階と、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令を分割せずに送信し、その応答の属性値を確認して、前記つなぎ目であることを示す属性値がある場合は前記ファイルシステム補完手段に処理をあげるつなぎ目処理判断手段とをさらに備える請求項7、12、14のいずれか1項記載のスイッチ方法。

【請求項16】 配下に、前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目を含むオブジェクトの移動要求があった場合に、前記つなぎ目以下のオブジェクトの実際の移動を行わず、当該つなぎ目以下のオブジェクトを所望の位置へと移動したと同等にクライアントに見せるよう、前記第4のテーブルの情報を書き換える段階をさらに備える請求項7、12、14、15のいずれか1項記載のスイッチ方法。

【請求項17】 前記サーバが公開しているオブジェクト群を別のサーバへ移動する処理を行う際、移動する全てのオブジェクトの移動処理の進行状態と、クライアントから移動処理中に要求された更新情報と、移動前における当該オブジェクトの移動元オブジェクトIDと、移動後における当該オブジェクトの移動先オブジェクトIDと、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管理する第6のテーブルを参照し、

前記疑似ファイルシステムを参照し、所定の基準に基づいて前記オブジェクトを移動させるとともに、移動処理の進行状態と、コピー処理中の更新ログと、当該オブジェクトの移動元IDと、移動先オブジェクトIDと、移動先サーバ識別情報とを前記第6のテーブルにエントリする段階と、

移動処理を実行している際に、クライアントからの移動対象オブジェクトへの要求を転送する方法を、前記第6のテーブルに登録されている進行状態と前記要求の内容により転送先を変更し、移動処理中であることを隠蔽する段階と、をさらに備える請求項6乃至8のいずれか1項に記載のスイッチ方法。

【請求項18】 前記移動処理中のオブジェクト群に対し、移動先へのデータのコピーが終了した前記オブジェクト群の一部オブジェクトの前記第6のテーブルの移動処理の進行状態の情報を移動終了と共に、前記疑似ファイルシステムにおけるマッピング情報を、移動元のオブジェクトから移動先のオブジェクトに変更する段階を備える請求項17記載のスイッチ方法。

【請求項19】 移動先オブジェクトにマッピングが変更され移動終了となったオブジェクトに対し、移動元オ

プロジェクトに関連付けられたオブジェクトIDによる要求であるかを、前記第6のテーブルにより判別し、該当する移動元オブジェクトIDを移動先オブジェクトIDへ書き換えて転送する段階を備える、請求項18記載のスイッチ方法。

【請求項20】 前記移動処理中にコピーが終了したオブジェクトに対し、クライアントからデータの書き込み要求がきた場合、移動元、および移動先へ書き込み要求を転送し、同期を維持することを特徴とする請求項17乃至19いずれか1項記載のスイッチ方法。

【請求項21】 前記移動処理中にコピー処理中のオブジェクトに対し、クライアントからのデータの書き込み要求が来た場合、書き込みデータの位置を前記第6のテーブルに更新情報として登録した後、書き込み要求を移動元に転送し、前記コピー処理終了後に、前記更新情報を参照して、移動元から更新部分のデータを読み出し、移動先のオブジェクトへ書き込みを行う請求項17乃至19のいずれか1項記載のスイッチ方法。

【請求項22】 前記移動処理中にコピー処理中のオブジェクトに対し、クライアントからのデータの書き込み要求が来た場合、書き込み要求に含まれる更新データおよび、書き込みデータの位置を第6のテーブルに更新情報として登録した後、書き込み要求を移動元へ転送し、前記コピー処理終了後に、前記更新情報を用いて、移動先のオブジェクトへ書き込みを行う請求項18乃至19のいずれか1項記載のスイッチ方法。

【請求項23】 少なくとも1つのクライアントと少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置されたスイッチ装置における、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ方法において、

前記オブジェクトIDと、当該元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報との組み合わせを管理する第1のテーブルを参照して、要求に含まれるオブジェクトIDに対応するサーバ識別情報を求める段階と、前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理する第2のテーブルを参照して、当該サーバ識別情報に対応するサーバのアドレス情報を求める段階と、

当該アドレス情報に基づいて当該要求の転送先を決定する段階と、を備えることを特徴とするスイッチ方法。

【請求項24】 前記第1のテーブルにおいて、既にエントリされているオブジェクトIDと同一のデータ列を有するオブジェクトIDをエントリする場合には、オブジェクトIDのデータ列を別のデータ列に書き換え、前記第1のテーブルに書き換えられたデータ列をエントリ

する請求項23記載のスイッチ方法。

【請求項25】 前記オブジェクトIDを管理するためのテーブルを、使用するプロトコル、およびそのバージョンに対応させて複数使用することを特徴とする請求項1乃至18、23、24のいずれか1項記載のスイッチ方法。

【請求項26】 少なくとも1つのクライアントと、少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置され、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ装置において、転送する応答に、前記オブジェクトを識別するために前記サーバで生成された元オブジェクトIDが含まれている場合には、前記元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報を挿入することによって前記元オブジェクトIDを情報付帯オブジェクトIDに書き換え、転送する要求に、前記情報付帯オブジェクトIDが含まれている場合には、前記情報付帯オブジェクトIDを前記元オブジェクトIDに復元するオブジェクトID書き換え手段と、

前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理するテーブルと、当該テーブルを参照して、転送する要求の情報付帯オブジェクトIDに含まれる前記サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する転送手段とを備えることを特徴とするスイッチ装置。

【請求項27】 前記サーバ識別情報には、オブジェクトが公開される単位である共有オブジェクト単位を識別するための共有オブジェクト単位識別情報が含まれる請求項26記載のスイッチ装置。

【請求項28】 前記要求および前記応答を関連づけ、前記要求および前記応答を前記クライアントおよび前記サーバで区別するために付与されているトランザクションIDとクライアントのアドレス情報との組み合わせを管理する第2のテーブルと、転送する応答に含まれるトランザクションIDに対応するアドレス情報を、前記第2のテーブルを参照して求め、当該アドレス情報に基づいて当該応答の転送先を決定する第2の転送手段とをさらに備える請求項26または27記載のスイッチ装置。

【請求項29】 前記要求および前記応答を関連づけ、前記クライアントと前記サーバの組み合わせを一意に特定するのに利用可能なOSモデル第4層以下のタグ情報とクライアントのアドレス情報との組み合わせを管理する分類テーブルと、
50 応答を転送すべきクライアントのアドレス情報を、前記

7

タグ情報と前記分類テーブルを参照して求め、当該アドレス情報に基づいて当該応答の転送先を決定する第3の転送手段とをさらに備える請求項26または27記載のスイッチ装置。

【請求項30】 前記要求および前記応答を関連づけ、前記クライアントと前記サーバの組み合わせを一意に特定するのに利用可能なサーバとの通信セッション・ポート番号とクライアントのアドレス情報との組み合わせを管理する分類テーブルと、

応答を転送すべきクライアントのアドレス情報を、前記ポート番号と前記分類テーブルを参照して求め、当該アドレス情報に基づいて当該応答の転送先を決定する第3の転送手段とをさらに備える請求項26または27記載のスイッチ装置。

【請求項31】 前記オブジェクトID書き換え手段は、前記元オブジェクトIDを前記情報付帯オブジェクトに書き換える際のアルゴリズムを識別するためのアルゴリズム識別情報と前記元オブジェクトIDに復元するための復元情報との組み合わせを管理する第3のテーブルと、

前記サーバ識別情報および前記アルゴリズム識別情報を前記元オブジェクトIDに挿入して前記情報付帯オブジェクトIDに書き換えるとともに、前記アルゴリズム識別情報と前記復元情報とを前記第3のテーブルに登録する応答処理手段と、

前記情報付帯オブジェクトIDに含まれるアルゴリズム識別情報に対応する復元情報を前記第3のテーブルを参照して求め、当該復元情報に基づいて前記情報付帯オブジェクトIDを前記元オブジェクトIDに復元する要求処理手段とを備える請求項26乃至30のいずれか1項記載のスイッチ装置。

【請求項32】 前記復元情報には、前記元オブジェクトIDから削除されたデータと、該データの位置が含まれる請求項31記載のスイッチ装置。

【請求項33】 前記応答処理手段は、前記元オブジェクトIDのデータ長が固定長である場合には、前記元オブジェクトIDの一部を削除し、削除した位置に前記アルゴリズム識別情報と前記サーバ識別情報とを挿入することによって前記情報付帯オブジェクトIDを生成する請求項32記載のスイッチ装置。

【請求項34】 前記応答処理手段は、前記元オブジェクトIDのデータ長が可変長であって、前記元オブジェクトIDと前記アルゴリズム識別情報と前記サーバ識別情報とのデータ長の和が最大データ長を越えない場合には、前記元オブジェクトIDに前記アルゴリズム識別情報と前記サーバ識別情報とを追加することによって前記情報付帯オブジェクトIDを生成し、前記元オブジェクトIDと前記アルゴリズム識別情報と前記サーバ識別情報とのデータ長の和が最大データ長を

8

越える場合には、前記元オブジェクトIDの一部を削除し、削除した位置に前記アルゴリズム識別情報と前記サーバ識別情報とを挿入することによって前記情報付帯オブジェクトIDを生成する請求項32または33記載のスイッチ装置。

【請求項35】 全ての要求および応答に含まれる情報付帯オブジェクトIDにおける前記アルゴリズム識別情報および前記サーバ識別情報の挿入位置が同一である請求項32乃至34のいずれか1項記載のスイッチ装置。

【請求項36】 全ての要求および応答に含まれる情報付帯オブジェクトIDにおける前記アルゴリズム識別情報および前記サーバ識別情報のデータ長が同一である請求項35項記載のスイッチ装置。

【請求項37】 前記サーバより公開されたオブジェクトのディレクトリツリーを組み合わせ、1つの統合ディレクトリツリーとして管理する擬似ファイルシステムをさらに備える請求項26乃至36のいずれか1項記載のスイッチ装置。

【請求項38】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令情報との組み合わせを管理する第4のテーブルと、

転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を、前記各送信先に転送する要求に分割し、前記各送信先から受信した全ての応答を、前記クライアントに転送する応答として合成するファイルシステム補完手段とをさらに備える請求項37記載のスイッチ装置。

【請求項39】 前記第4のテーブルに登録されるオブジェクトIDは、前記情報付帯オブジェクトIDである請求項38記載のスイッチ装置。

【請求項40】 前記サーバが公開するオブジェクトが移動したときの、移動前における当該オブジェクトの移動元オブジェクトIDと、移動後における当該オブジェクトの移動先オブジェクトIDと、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管理する第5のテーブルと、

前記擬似ファイルシステムによって管理されている統合ディレクトリツリーを参照し、所定の基準に基づいて前記オブジェクトを移動させるとともに、当該オブジェクトの移動元オブジェクトIDと、移動先オブジェクトIDと、移動先サーバ識別情報とを前記第5のテーブルにエントリするデータ移動処理手段と、転送する要求に含まれるオブジェクトIDが前記第5のテーブルの移動元オブジェクトIDと一致する場合には、当該オブジェクトIDを、一致した当該移動元オブジェクトIDに対応する移動先オブジェクトIDに変更し、当該移動元オブ

ジェクトIDに対応する移動先サーバ識別情報に基づいて送信先のサーバを決定する移動隠蔽手段とをさらに備える請求項37乃至39のいずれか1項記載のスイッチ装置。

【請求項41】 前記第5のテーブルには、移動元オブジェクトIDとして、前記情報付帯オブジェクトIDがエントリされる請求項40記載のスイッチ装置。

【請求項42】 前記オブジェクトID書き換え手段は、情報付帯オブジェクトIDに、データが移動したことを示すデータ移動フラグを組み込み、転送する要求に含まれるオブジェクトIDが、前記第5のテーブルの移動元オブジェクトIDと一致する場合には、前記データ移動フラグを真に設定し、

前記移動隠蔽手段は、前記データ移動フラグが真である場合に限って、前記第5のテーブルを検索する請求項40または41記載のスイッチ装置。

【請求項43】 前記第5のテーブルでは、1回のデータ移動を1世代として付与される世代番号ごとにテーブルが作成され、

前記オブジェクトID書き換え手段は、転送する要求に含まれるオブジェクトIDが前記第5のテーブルの移動元オブジェクトIDと一致する場合には、前記情報付帯オブジェクトIDに前記世代番号を組み込み、

前記移動隠蔽手段は、前記第5のテーブルにおける各テーブルのうち、前記世代番号が付与されたテーブルだけを検索する請求項40乃至42のいずれか1項記載のスイッチ装置。

【請求項44】 前記統合ディレクトリツリー管理において特殊操作の必要な命令情報との組み合わせを管理する第4のテーブルと、

転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を他の要求とは異なる通信セッションポート（ポート番号X）を用いて送信し、前記各送信先から受信した応答を、その他の応答から当該ポート番号Xにより識別し、前記クライアントに転送する応答として復路の特殊処理を行うファイルシステム補完手段とをさらに備える請求項37記載のスイッチ装置。

【請求項45】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令情報との組み合わせを管理する第4のテーブルと、

転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を、前記各送信先に転送する要求に分割し、当該分割された要求を他の要求とは異なる通信セッションポートに対応するポート番号を用いて送信し、前記各送信先から受信した全ての応答を、その他

の応答から当該ポート番号X、Y・・・により識別し、前記クライアントに転送する応答として合成するファイルシステム補完手段とをさらに備える請求項37記載のスイッチ装置。

【請求項46】 前記統合ディレクトリツリー管理において特殊操作の必要な命令情報との組み合わせを管理する第4のテーブルと、

特殊操作を必要とする要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、OS Iモデル第4層以下で判断できるタグ情報を当該要求に付与してこれを送信し、前記各送信先から受信した応答を、当該タグ情報により識別し、前記クライアントに転送する応答として復路の特殊処理を行うファイルシステム補完手段とをさらに備える請求項37記載のスイッチ装置。

【請求項47】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと、当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令情報との組み合わせを管理する第4のテーブルと、

転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を前記各送信先に転送する要求に分割し、当該分割された要求にOS Iモデル第4層以下で判断できるタグ情報を付与してこれらを送信し、前記各送信先から受信した全ての応答を、当該タグ情報により識別し、前記クライアントに転送する応答として合成するファイルシステム補完手段とをさらに備える請求項37記載のスイッチ装置。

【請求項48】 前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトの属性に、当該オブジェクトがつなぎ目であることを示す属性を記録する擬似ファイルシステム管理手段と、

当該オブジェクトIDに関連し複数の送信先へのアクセスを必要とする命令を分割せずに送信し、その応答の属性値を確認して、前記つなぎ目であることを示す属性値がある場合は前記ファイルシステム補完手段に処理をあげるつなぎ目処理判断手段とをさらに備える請求項38、45、47のいずれか1項記載のスイッチ装置。

【請求項49】 前記サーバより公開されたオブジェクトのディレクトリツリーを組み合わせ、1つの統合ディレクトリツリーとして管理する請求項35記載の擬似ファイルシステムに、他の種類のOSまたはプロトコルで管理される統合ディレクトリツリーの情報を読みディレクトリツリーファイルシステムに反映させる融合エージェントとをさらに備える請求項37記載のスイッチ装置。

【請求項50】 配下に、前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリー

のつなぎ目を含むオブジェクトの移動要求があった場合に、前記つなぎ目以下のオブジェクトの実際の移動を行わず、当該つなぎ目以下のオブジェクトを所望の位置へと移動したと同等にクライアントに見せるよう請求項1記載のつなぎ目の管理を行うファイルシステム補完手段の第4のテーブルの情報を書き換える高速オブジェクト移動手段を更に備える請求項38、45、47、48のいずれか1項記載のスイッチ装置。

【請求項51】 前記サーバが公開しているオブジェクト群を別のサーバへ移動する処理を行う際、移動する全てのオブジェクトの移動処理の進行状態と、クライアントから移動処理中に要求された更新情報と、移動前における当該オブジェクトの移動元オブジェクトIDと、移動後における当該オブジェクトの移動先オブジェクトIDと、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管理する第6のテーブルと、前記疑似ファイルシステムを参照し、所定の基準に基づいて前記オブジェクトを移動させるとともに、移動処理の進行状態と、コピー処理中の更新ログと、当該オブジェクトの移動元IDと、移動先オブジェクトIDと、移動先サーバ識別情報とを前記第6のテーブルにエントリする移動処理管理手段と、を備え、移動処理を実行している際に、クライアントからの移動対象オブジェクトへの要求を転送する方法を、前記第6のテーブルに登録されている進行状態と前記要求の内容により転送先を変更して、移動処理中であることが隠蔽されるようにする、請求項37乃至39のいずれか1項記載のスイッチ装置。

【請求項52】 前記移動処理中のオブジェクト群に対し、移動先へのデータのコピーが終了した前記オブジェクト群の一部オブジェクトの前記第6のテーブルの移動処理の進行状態の情報を移動終了と共に、前記疑似ファイルシステムにおけるマッピング情報を、移動元のオブジェクトから移動先のオブジェクトに変更する、マッピング情報変更手段を備える請求項51記載のスイッチ装置。

【請求項53】 前記マッピング情報変更手段により、移動先オブジェクトにマッピングが変更され、移動終了となったオブジェクトに対し、移動元オブジェクトに関連付けられたオブジェクトIDによる要求であるかを、前記第6のテーブルにより判別し、該当する移動元オブジェクトIDを移動先オブジェクトIDへ書き換えて転送する、オブジェクトID書き換え手段を備える、請求項52記載のスイッチ装置。

【請求項54】 前記移動処理中にコピーが終了したオブジェクトに対し、クライアントからデータの書き込み要求がきた場合、移動元、および移動先へ書き込み要求を転送し、同期を維持することを特徴とする請求項51乃至53いずれか1項記載のスイッチ装置。

【請求項55】 前記移動処理中にコピー処理中のオブ

ジェクトに対し、クライアントからのデータの書き込み要求が来た場合、書き込みデータの位置を前記第6のテーブルに更新情報として登録した後、書き込み要求を移動元に転送し、前記コピー処理終了後に、前記更新情報を参照して、移動元から更新部分のデータを読み出し、移動先のオブジェクトへ書き込みを行う請求項51乃至53のいずれか1項記載のスイッチ装置。

【請求項56】 前記移動処理中にコピー処理中のオブジェクトに対し、クライアントからのデータの書き込み要求が来た場合、書き込み要求に含まれる更新データおよび、書き込みデータの位置を第6のテーブルに更新情報として登録した後、書き込み要求を移動元へ転送し、前記コピー処理終了後に、前記更新情報を用いて、移動先のオブジェクトへ書き込みを行う請求項51乃至53のいずれか1項記載のスイッチ装置。

【請求項57】 少なくとも1つのクライアントと、少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置され、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ装置において、前記オブジェクトIDと、当該元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報との組み合わせを管理する第1のテーブルと、

前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理する第2のテーブルと、前記第1のテーブルを参照して、要求に含まれるオブジェクトIDに対応するサーバ識別情報を求め、前記第2のテーブルを参照して、当該サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する転送手段とを備えることを特徴とするスイッチ装置。

【請求項58】 前記第1のテーブルにおいて、既にエントリされているオブジェクトIDと同一のデータ列を有するオブジェクトIDをエントリする場合には、オブジェクトIDのデータ列を別のデータ列に書き換え、前記第1のテーブルに書き換えられたデータ列をエントリする請求項57記載のスイッチ装置。

【請求項59】 前記オブジェクトIDを管理するためのテーブルを、使用するプロトコル、およびそのバージョンに対応させて複数設けることを特徴とする請求項26乃至52、57、58のいずれか1項記載のスイッチ装置。

【請求項60】 少なくとも1つのクライアントと、少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置されたコンピュータに、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに

転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送する処理を実行させるためのプログラムにおいて、

転送する応答に、前記オブジェクトを識別するために前記サーバで生成された元オブジェクト ID が含まれている場合には、前記元オブジェクト ID に対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報を挿入することによって前記元オブジェクト ID を情報付帯オブジェクト ID に書き換え、転送する要求に、前記情報付帯オブジェクト ID が含まれている場合には、前記情報付帯オブジェクト ID を前記元オブジェクト ID に復元し、

前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理するテーブルを参照し、転送する要求の情報付帯オブジェクト ID に含まれる前記サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する処理を前記コンピュータに実行させるためのプログラム。

【請求項 6 1】 前記要求および前記応答に関連づけ、前記要求および前記応答を前記クライアントおよび前記サーバで区別するために付与されているトランザクション ID とクライアントのアドレス情報との組み合わせを管理する第 2 のテーブルを参照して、転送する応答に含まれるトランザクション ID に対応するアドレス情報を求め、当該アドレス情報に基づいて当該応答の転送先を決定する処理を前記コンピュータにさらに実行させる請求項 6 0 記載のプログラム。

【請求項 6 2】 前記サーバより公開されたオブジェクトのディレクトリツリーを組み合わせ、1 つの統合ディレクトリツリーとして管理する擬似ファイルシステムを備えておき、

前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクト ID と当該オブジェクト ID に関連し、複数の送信先に渡るアクセスを必要とする命令情報との組み合わせを管理する第 4 のテーブルを参照して、転送する要求に含まれるオブジェクト ID と命令情報との組み合わせが、前記第 4 のテーブルにエントリされている場合には、当該要求を、前記各送信先に転送する要求に分割し、

前記各送信先から受信した全ての応答を、前記クライアントに転送する応答に合成する処理を前記コンピュータにさらに実行させる請求項 6 1 または 6 2 記載のプログラム。

【請求項 6 3】 前記サーバが公開するオブジェクトが移動したときの、移動前における当該オブジェクトの移動元オブジェクト ID と、移動後における当該オブジェクトの移動先オブジェクト ID と、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管

理する第 5 のテーブルを参照し、所定の基準に基づいて前記オブジェクトを前記擬似ファイルシステムによって管理されている統合ディレクトリツリーを参照して移動させるとともに、当該オブジェクトの移動元オブジェクト ID と、移動先オブジェクト ID と、移動先サーバ識別情報とを前記第 5 のテーブルにエントリし、

転送する要求に含まれるオブジェクト ID が前記第 5 のテーブルの移動元オブジェクト ID と一致する場合には、当該オブジェクト ID を当該移動元オブジェクト ID に対応する移動先オブジェクト ID に変更し、当該移動元オブジェクト ID に対応する移動先サーバ識別情報に基づいて送信先のサーバを決定する処理を前記コンピュータにさらに実行させる請求項 6 2 記載のプログラム。

【請求項 6 4】 少なくとも 1 つのクライアントと、少なくとも 1 つのオブジェクトを管理する少なくとも 1 つのサーバとの間に論理的に配置されたコンピュータに、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送する処理を実行させるためのプログラムにおいて、

前記オブジェクトと、当該元オブジェクト ID に対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報との組み合わせを管理する第 1 のテーブルを参照し、前記要求に含まれるオブジェクト ID に対応するサーバ識別情報を求め、

前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理する第 2 のテーブルを参照して、当該サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する処理を前記コンピュータに実行させることを特徴とするプログラム。

【請求項 6 5】 前記第 1 のテーブルにおいて、既にエントリされているオブジェクト ID と同一のデータ列を有するオブジェクト ID をエントリする場合には、オブジェクト ID のデータ列を別のデータ列に書き換え、前記第 1 のテーブルに書き換えられたデータ列をエントリしておく処理を前記コンピュータにさらに実行させる請求項 6 4 記載のプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、クライアントとサーバとの間に論理的に配置されるスイッチ装置に関し、特に、ネットワークを介してファイル処理を実行するリモートファイルシステムが実装されたクライアント・サーバ・システムにおいて、クライアントに透過なシングルシステムイメージ (SSI) の提供や負荷分散などのネットワークトラフィック制御を行うスイッチ装置に関する。

【0002】

【従来の技術】ネットワークを介したクライアント・サーバ・システムを実現する方法としては、RFC (request for comments) 1094やRFC 1813に記載されたネットワークファイルシステム (NFS; network file system) プロトコルを使用する方法が有名である。このNFSプロトコルは、NFSを実装するクライアントに、ネットワークを介した遠隔地にあるサーバに対するファイルの作成や消去、読み書き、移動等の、ファイルシステムとしての基本的な動作や、他のクライアントとのファイル共有などの利便性を提供する。NFSプロトコルでは、クライアントは、サーバ上のファイルなどのオブジェクトを識別するためのオブジェクトIDを使用し、リモートプロシージャコール (RPC) などのアプリケーションインターフェースを介してサーバにアクセスを行う。例えば、NFSでは、オブジェクトIDとして、ファイルハンドルとよばれるものがある。

【0003】(レイヤ4-7スイッチ) ネットワークファイルシステムの実装は、いわゆるIP (Internet Protocol) ネットワークのOSI (開放型システム間相互接続; open systems interconnection) 参照モデルに準拠したものも多く、OSIモデルのレイヤ3まで(物理層、データリンク層、ネットワーク層)の情報を利用して通常のルータなどの経路制御が行われることは広く知られている。さらに、近年では、OSI参照モデルのレイヤ4からレイヤ7まで(トランスポート層、セッション層、プレゼンテーション層、アプリケーション層)の情報を利用して、さらにさまざまな利便性を提供しようという動きが盛んになっている。

【0004】レイヤ4情報を利用して送られてくるパケットのプロトコルの種類を判別し、プロトコルごとに対応するサーバへ振り分けるネットワークスイッチ装置は、ウェブ(web)ホスティングサービスなどで広く利用されている。さらに、最近では、レイヤ7情報を利用したウェブコンテンツごとのサーバ振り分けを実現するネットワークスイッチ装置が製品として発表されている。これらの技術の主な目的は、サーバの負荷分散や、サーバへのアクセスポリシーの制御(例えば、利用者の認証を行うSSL (secure socket layer) 処理の単一サーバでの実行、年齢別の閲覧制限などである。SSL処理を行う場合には、負荷分散等の目的のためのサーバの切り換えは禁止される。)である。なお、スイッチ装置は、論理的なパケット・スイッチとして動作する。

【0005】図42は、ウェブスイッチの機能を示すブロック図である。ウェブスイッチ900は、パケット処理部903と、HTTP (hypertext transfer protocol) 処理部902とを備えている。パケット処理部903で受信したパケット処理を行った後、HTTP処理部902は、レイヤ4からレイヤ7に関する情報の収集を行い、レイヤ7に分類されるURL (uniform resource lo

cator)に関する情報を取り出す。さらに、HTTP処理部902は、取り出したURLに関する情報に基づいて、このURLを扱うべきサーバを、その所在はもとより、負荷分散などのポリシーも考慮して検索、特定する。そして、HTTP処理部902は、そのサーバへパケットが転送されるように、グローバルアドレスからローカルアドレスの変換 (NAT (network address translator) 型変換) や、後述する仮想IPクラスタリングによるMAC (media access control) アドレスの書き換えを行うことによって、元パケットのMACアドレスやIPアドレスを変換し、その変換に伴うチェックサムなどのセキュリティの矛盾を訂正する。パケット処理部903は、書き換えおよび訂正されたパケットをネットワーク上に送出する。

【0006】また、上述したスイッチ装置の以外にも、負荷分散等の目的のためのさまざまな技術が公開されている。そのような技術として広く知られるものに、例えば、各ネットワークエリアに分散配置した各DNS (ドメインネームサーバ) が1つのドメイン名に複数のIPアドレスを対応させ、ラウンドロビン方式でIPアドレスを選択してクライアントへ伝えるという方法がある。

【0007】(仮想IPクラスタリング) ところで、仮想IPアドレスをクラスタアドレスとして定義し、複数のサーバにIPエイリアスとしてこの仮想IPアドレスを設定し、スイッチとしての役割を有するコーディネータが、MACアドレスの書き換えによりパケットを適切なサーバへリダイレクトして負荷分散を行う方法が知られている。このような方法では、各サーバが、仮想IPアドレスをIPエイリアスとして設定しているので、クライアントからはあたかも1つのサーバが応答したように見える。この方法では、コーディネータを介さずに、直接対応したサーバからクライアントへ応答を返すことができる。応答を直接返す方法は、パケットがコーディネータを復路も通過する方法と比べ、ボトルネックが生じにくくなるという利点がある。この方法、およびこの方法に若干の改良を加えた方法は、現在、リナックス (Linux) などのOS (オペレーティングシステム) を用いたクラスタリングシステムの一構成法として広く普及しているものである。

【0008】また、特開2001-51890号公報には、LAN (ローカルエリアネットワーク; local area network) などのネットワーク上に分散した複数のファイルサーバの台数やストレージ装置の接続状態をクライアントに意識させないで済むようにする別の方法が開示されている。図43に示すように、この方法を実行するクライアント・サーバ・システムでは、ネットワーク2上に分散して構築され、互いにサーバ間専用結線7によって接続され、マルチキャストアドレスを共有する複数のファイルサーバ3を備えている。そして、各ファイルサーバ3にまたがって動作する仮想分散ファイルシステ

ム4が実装されている。このシステムを構成する、各ファイルサーバ3上のモジュールは、クライアント1からマルチキャストされたファイル操作要求を受信すると、仮想分散ファイルシステム4と全ローカルファイルシステム6とのマッピングテーブルを利用して、自サーバが上記要求を処理可能な最適なサーバであるか否かを判断し、その判断結果に基づいて要求されたファイル操作を他のファイルサーバ3と重複することなく処理する。こうすることによって、ネットワーク上に分散した複数のファイルサーバ3を、クライアント1からは単一のサーバとして扱うことができるようになるため、サーバ3の台数やストレージ装置の接続状態をクライアント1に意識させない仮想分散ファイルサーバシステムを提供することができる。

【0009】

【発明が解決しようとする課題】しかし、分散型のネットワークを介したファイルシステムサービスに上述した従来技術を適用した場合には、以下に示す問題点が発生する。

【0010】(1) レイヤ4-7スイッチを用いる方法も、仮想IPクラスタリングを用いる方法も、パケットのリダイレクションにあたって書き換える内容がMACアドレスもしくはIPアドレスに限定されている。そのため、NFSのファイルハンドルなどのオブジェクトIDをクライアントがキャッシュに保管する実装のプロトコルにおいて、管理者が負荷分散や容量分散の目的という本来クライアントに対して通知する必要のない操作を行う場合、例えば、クライアントに一度割り当てたサーバから別のサーバへオブジェクトを移動するような場合には、クライアントが、古いサーバのオブジェクトIDをパケット内に含めてサーバにアクセスしてしまう。したがって、スイッチが新しいサーバにそのパケットを転送しても、新しいサーバはそのパケット内のオブジェクトIDを判読できずに、クライアントにエラーを返してしまう。すなわち、負荷分散型ネットワークを介したファイルシステムサービスにおいては、従来の負荷分散の方法を適用しても、クライアントに対する透過性を実現することができない。

【0011】(2) IPエイリアスの設定や仮想分散ファイルシステムの導入を行う際には、ユーザがすでに利用しているファイルサーバに対する改造が必須となるので、既存のクライアント・サーバのファイルシステムが稼動している場合にその構成を活かしたまま提供される利便性を得るためには、負荷分散という目的のみを解決する場合でも、煩雑なインストール作業が発生してしまう。

【0012】本発明の目的は、クライアントおよびサーバが標準のネットワークを介したファイルシステムサービスプロトコル(例えば、NFS)さえサポートしていれば、それらの既存システムに手を加えることなく、複

数のサーバを統合管理してクライアントにシングルシステムイメージ(SSI; single system image)のファイルサービスを提供し、負荷分散、容量割り当て等の管理目的によるオブジェクトの再配置をユーザから隠蔽してSSIを維持するなどの利便性を実現することができるスイッチ方法及び装置を提供することである。

【0013】

【課題を解決するための手段】上記課題を解決するために、本発明のスイッチ方法は、少なくとも1つのクライアントと少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置されたスイッチ装置における、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ方法において、転送する応答に、前記オブジェクトを識別するために前記サーバで生成された元オブジェクトIDが含まれている場合には、前記元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報を挿入することによって前記元オブジェクトIDを情報付帯オブジェクトIDに書き換える段階と、転送する要求に、前記情報付帯オブジェクトIDが含まれている場合には、前記情報付帯オブジェクトIDを前記元オブジェクトIDに復元する段階と、前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理するテーブルを参照して、転送する要求の情報付帯オブジェクトIDに含まれる前記サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する段階と、を備えることを特徴とする。

【0014】また、本発明の別のスイッチ方法は、少なくとも1つのクライアントと少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置されたスイッチ装置における、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ方法において、前記オブジェクトIDと、当該元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報との組み合わせを管理する第1のテーブルを参照して、要求に含まれるオブジェクトIDに対応するサーバ識別情報を求める段階と、前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理する第2のテーブルを参照して、当該サーバ識別情報に対応するサーバのアドレス情報を求める段階と、当該アドレス情報に基づいて当該要求の転送先を決定する段階と、を備えることを特徴とする。

【0015】本発明のスイッチ装置は、少なくとも1つ

のクライアントと、少なくとも1つのオブジェクトを管理する少なくとも1つのサーバとの間に論理的に配置され、前記クライアントから送信された前記オブジェクトに関する要求を、当該オブジェクトを管理する前記サーバに転送し、当該要求に対する前記サーバからの応答を前記要求の送信元であるクライアントに転送するスイッチ装置において、転送する応答に、前記オブジェクトを識別するために前記サーバで生成された元オブジェクトIDが含まれている場合には、前記元オブジェクトIDに対応するオブジェクトを管理するサーバを他のサーバから識別するためのサーバ識別情報を挿入することによって前記元オブジェクトIDを情報付帯オブジェクトIDに書き換え、転送する要求に、前記情報付帯オブジェクトIDが含まれている場合には、前記情報付帯オブジェクトIDを前記元オブジェクトIDに復元するオブジェクトID書き換え手段と、前記サーバのアドレス情報と前記サーバ識別情報との組み合わせを管理するテーブルと、該テーブルを参照して、転送する要求の情報付帯オブジェクトIDに含まれる前記サーバ識別情報に対応するサーバのアドレス情報を求め、当該アドレス情報に基づいて当該要求の転送先を決定する転送手段とを備えることを特徴とする。

【0016】本発明のスイッチ装置では、オブジェクトを管理するサーバとそのオブジェクトのオブジェクトIDとを対応付けすることができるスイッチを備え、そのスイッチが、クライアントに送信するオブジェクトIDに、そのオブジェクトを管理するサーバの識別情報を付与する。このようにすれば、次にクライアントがそのオブジェクトにアクセスしようとするときにクライアントが送信する要求中のオブジェクトIDに、そのオブジェクトを管理するサーバ識別情報が含まれるようになるため、スイッチ装置は、要求に含まれるサーバ識別情報から、転送先のサーバを決定することができるようになる。したがって、本発明のスイッチ装置を適用すれば、クライアント側で、オブジェクトを管理するサーバを特定する必要がなくなるため、クライアントに対する透過性を実現することができる。

【0017】また、本発明のスイッチ装置を適用しても、クライアントおよびサーバの処理を変更する必要はない。したがって、クライアントおよびサーバが標準のネットワークを介したファイルシステムサービスプロトコルさえサポートしていれば、既存システムに手を加えることなく、複数のサーバを統合管理してクライアントにSSIのファイルサービスを提供することができる。

【0018】本発明のスイッチ装置では、前記サーバ識別情報には、オブジェクトが公開される単位である共有オブジェクト単位を識別するための共有オブジェクト単位識別情報が含まれる。このようにすれば、サーバ単位よりもさらに分解能が高い共有オブジェクト単位でのきめ細かい転送管理を実現することが可能となる。

【0019】また、本発明のスイッチ装置では、前記要求および前記応答を関連づけ、前記要求および前記応答を前記クライアントおよび前記サーバで区別するために付与されているトランザクションIDとクライアントのアドレス情報との組み合わせを管理する第2のテーブルと、転送する応答に含まれるトランザクションIDに対応するアドレス情報を、前記第2のテーブルを参照して求め、当該アドレス情報に基づいて当該応答の転送先を決定する第2の転送手段とをさらに備えていてもよい。

【0020】また、本発明では、前記要求および前記応答を関連づけ、前記クライアントと前記サーバの組み合わせを一意に特定するのに利用可能なOS Iモデル第4層以下のタグ情報とクライアントのアドレス情報との組み合わせを管理する分類テーブルと、応答を転送すべきクライアントのアドレス情報を、前記タグ情報と前記分類テーブルを参照して求め、当該アドレス情報に基づいて当該応答の転送先を決定する第3の転送手段とをさらに設けてもよく、あるいは、前記要求および前記応答を関連づけ、前記クライアントと前記サーバの組み合わせを一意に特定するのに利用可能なサーバとの通信セッション・ポート番号とクライアントのアドレス情報との組み合わせを管理する分類テーブルと、応答を転送すべきクライアントのアドレス情報を、前記ポート番号と前記分類テーブルを参照して求め、当該アドレス情報に基づいて当該応答の転送先を決定する第3の転送手段とをさらに設けてもよい。

【0021】このように分類テーブルと第3の転送手段とを設けることにより、ハードウェア処理が可能なOS Iモデル第4層以下のタグ情報を用いることができるようになり、より高速に処理を行うことができるようになる。

【0022】本発明では、前記オブジェクトID書き換え手段は、前記元オブジェクトIDを前記情報付帯オブジェクトに書き換える際のアルゴリズムを識別するためのアルゴリズム識別情報と前記元オブジェクトIDに復元するための復元情報との組み合わせを管理する第3のテーブルと、前記サーバ識別情報および前記アルゴリズム識別情報を前記元オブジェクトIDに挿入して前記情報付帯オブジェクトIDに書き換えるとともに、前記アルゴリズム識別情報と前記復元情報とを前記第3のテーブルに登録する応答処理手段と、前記情報付帯オブジェクトIDに含まれるアルゴリズム識別情報に対応する復元情報を前記第3のテーブルを参照して求め、当該復元情報に基づいて前記情報付帯オブジェクトIDを前記元オブジェクトIDに復元する要求処理手段とを備えていてもよく、前記復元情報は、前記元オブジェクトIDから削除されたデータと、該データの位置が含まれていてもよい。

【0023】さらに、前記応答処理手段は、前記元オブジェクトIDのデータ長が固定長である場合には、前記

元オブジェクトIDの一部を削除し、削除した位置に前記アルゴリズム識別情報と前記サーバ識別情報とを挿入することによって前記情報付帯オブジェクトIDを生成し、前記元オブジェクトIDのデータ長が可変長であって、前記元オブジェクトIDと前記アルゴリズム識別情報と前記サーバ識別情報とのデータ長の和が最大データ長を越えない場合には、前記元オブジェクトIDに前記アルゴリズム識別情報と前記サーバ識別情報とを追加することによって前記情報付帯オブジェクトIDを生成し、前記元オブジェクトIDと前記アルゴリズム識別情報と前記サーバ識別情報とのデータ長の和が最大データ長を越える場合には、前記元オブジェクトIDの一部を削除し、削除した位置に前記アルゴリズム識別情報と前記サーバ識別情報とを挿入することによって前記情報付帯オブジェクトIDを生成するようにしてもよい。

【0024】さらに、本発明のスイッチ装置は、全ての要求および応答に含まれる情報付帯オブジェクトIDにおける前記アルゴリズム識別情報および前記サーバ識別情報の挿入位置やデータ長が同一であるのが望ましい。このようにすれば、オブジェクトIDに含まれる情報を抽出および検索するのに要する時間を短縮することができる。

【0025】また、本発明の他のスイッチ装置では、前記サーバより公開されたオブジェクトのディレクトリツリーを組み合わせ、1つの統合ディレクトリツリーとして管理する擬似ファイルシステムをさらに備え、前記統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目に相当するオブジェクトのオブジェクトIDと当該オブジェクトIDに関連し、複数の送信先に渡るアクセスを必要とする命令情報との組み合わせを管理する第4のテーブルと、転送する要求に含まれるオブジェクトIDと命令情報との組み合わせが、前記第4のテーブルにエントリされている場合には、当該要求を、前記各送信先に転送する要求に分割し、前記各送信先から受信した全ての応答を、前記クライアントに転送する応答に合成するファイルシステム補完手段とをさらに備える。

【0026】本発明のスイッチ装置では、擬似ファイルシステムの統合ディレクトリツリーを参照して、複数のサーバあるいは複数のファイルシステムを結合しているツリー部分のみの名前解決処理を行い、その他のツリー部分についてはサーバに処理を任せる。このようにすれば、名前解決処理の負荷を、スイッチとサーバとの間で分散することができる。なお、検索処理の効率化から、前記第4のテーブルに登録されるオブジェクトIDは、要求に含まれるオブジェクトIDと同じ前記情報付帯オブジェクトIDであるのが望ましい。

【0027】また、本発明の他のスイッチ装置では、前記サーバが公開するオブジェクトが移動したときの、移動前における当該オブジェクトの移動元オブジェクトID

Dと、移動後における当該オブジェクトの移動先オブジェクトIDと、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管理する第5のテーブルと、所定の基準に基づいて前記オブジェクトを前記擬似ファイルシステムによって管理されている統合ディレクトリツリーを参照して移動させるとともに、当該オブジェクトの移動元オブジェクトIDと、移動先オブジェクトIDと、移動先サーバ識別情報とを前記第5のテーブルにエントリするデータ移動処理手段と、転送する要求に含まれるオブジェクトIDが前記第5のテーブルの移動元オブジェクトIDと一致する場合には、前記第5のテーブルを参照し、当該オブジェクトIDを当該移動元オブジェクトIDに対応する移動先オブジェクトIDに変更し、当該移動元オブジェクトIDに対応するサーバ識別情報に基づいて送信先のサーバを決定する移動隠蔽手段とをさらに備える。

【0028】本発明のスイッチ装置では、移動元オブジェクトID、移動先オブジェクトID、移動先サーバとを対応付け可能な移動隠蔽手段を備えているので、負荷分散、容量割り当て等の管理目的によるオブジェクトの再配置をユーザから隠蔽してSSIを維持するなどの利便性を実現することができる。

【0029】また、本発明のスイッチ装置では、前記第5のテーブルには、移動元オブジェクトIDとして、前記情報付帯オブジェクトIDがエントリされるのが処理の効率化の観点から見て望ましい。

【0030】また、前記オブジェクトID書き換え手段は、情報付帯オブジェクトIDに、データが移動したことを示すデータ移動フラグを組み込み、転送する要求に含まれるオブジェクトIDが前記第5のテーブルの移動元オブジェクトIDと一致する場合には、前記データ移動フラグを真に設定し、前記移動隠蔽手段は、前記データ移動フラグが真である場合に限り、前記第5のテーブルを検索するのが処理の効率化の観点から見て望ましい。

【0031】さらに、前記第5のテーブルでは、1回のデータ移動を1世代として付与される世代番号ごとにテーブルが作成され、前記オブジェクトID書き換え手段は、転送する要求に含まれるオブジェクトIDが前記第5のテーブルの移動元オブジェクトIDと一致する場合には、前記情報付帯オブジェクトIDに前記世代番号を組み込み、前記移動隠蔽手段は、前記第5のテーブルにおける各テーブルのうち、前記世代番号が付与されたテーブルだけを検索するのが処理の効率化の観点から見て望ましい。

【0032】本発明のさらに別のスイッチ装置は、前記サーバが公開しているオブジェクト群を別のサーバへ移動する処理を行う際、移動する全てのオブジェクトの移動処理の進行状態と、クライアントから移動処理中に要求された更新情報と、移動前における当該オブジェクト

10

20

30

40

50

の移動元オブジェクトIDと、移動後における当該オブジェクトの移動先オブジェクトIDと、移動先のサーバの識別情報である移動先サーバ識別情報との組み合わせを管理する第6のテーブルと、前記擬似ファイルシステムを参照し、所定の基準に基づいて前記オブジェクトを移動させるとともに、移動処理の進行状態と、コピー処理中の更新ログと、当該オブジェクトの移動元IDと、移動先オブジェクトIDと、移動先サーバ識別情報とを前記第6のテーブルにエントリする移動処理管理手段と、を備え、移動処理を実行している際に、クライアントからの移動対象オブジェクトへの要求を転送する方法を、前記第6のテーブルに登録されている進行状態と前記要求の内容により転送先を変更して、移動処理中であることが隠蔽されるようにする。

【0033】この構成を採用することにより、管理者によるサーバ間でのオブジェクト再配置作業時においても、ユーザへは再配置作業を隠蔽させることで、システム停止や、ファイルアクセスサービス停止を発生させずに管理作業を実行することを可能にする。

【0034】この場合、前記移動処理中のオブジェクト群に対し、移動先へのデータのコピーが終了した前記オブジェクト群の一部オブジェクトの前記第6のテーブルの移動処理の進行状態の情報を移動終了と共に、前記擬似ファイルシステムにおけるマッピング情報を、移動元のオブジェクトから移動先のオブジェクトに変更する、マッピング情報変更手段を備えるようにしてもよい。また、前記マッピング情報変更手段により、移動先オブジェクトにマッピングが変更され、移動終了となったオブジェクトに対し、移動元オブジェクトに関連付けられたオブジェクトIDによる要求であるかを、前記第6のテーブルにより判別し、該当する移動元オブジェクトIDを移動先オブジェクトIDへ書き換えて転送する、オブジェクトID書き換え手段を備えるようにしてもよい。

【0035】さらに、前記移動処理中にコピーが終了したオブジェクトに対し、クライアントからデータの書き込み要求がきた場合、移動元、および移動先へ書き込み要求を転送し、同期を維持するようにしてもよく、前記移動処理中にコピー処理中のオブジェクトに対し、クライアントからのデータの書き込み要求が来た場合、書き込みデータの位置を前記第6のテーブルに更新情報として登録した後、書き込み要求を移動元に転送し、前記コピー処理終了後に、前記更新情報を参照して、移動元から更新部分のデータを読み出し、移動先のオブジェクトへ書き込みを行うようにしてもよく、前記移動処理中にコピー処理中のオブジェクトに対し、クライアントからのデータの書き込み要求が来た場合、書き込み要求に含まれる更新データおよび、書き込みデータの位置を第6のテーブルに更新情報として登録した後、書き込み要求を移動元へ転送し、前記コピー処理終了後に、前記更新情報を用いて、移動先のオブジェクトへ書き込みを行う

ようにしてもよい。

【0036】本発明においてオブジェクトIDとは、典型的にはNFSにおけるファイルハンドルであるが、これに限定されるものではなく、ネットワークを介在することができる任意のファイルシステムにおいて、オブジェクトを識別するものであればいずれのものも含まれる。また、ファイルシステムとしても、NFSに限定されるものではなく、他のファイルシステムにも本発明は適用可能である。

【0037】

【発明の実施の形態】次に、本発明の好ましい実施の形態のスイッチ装置について図面を参照して詳細に説明する。

【0038】（第1の実施形態）まず、本発明の第1の実施形態のスイッチ装置について説明する。

【0039】図1は、本発明の第1の実施形態のスイッチ装置が設けられたクライアント・サーバ形式のリモートファイルシステムであるネットワークファイルシステムのブロック図である。図1に示すように、このネットワークファイルシステムでは、少なくとも1つのクライアント1と、少なくとも1つのサーバ3と、本実施形態のスイッチ装置の一例であるスイッチ100とが、LANあるいはインターネット等であるネットワーク2に接続されている。各サーバ3は、ファイルアクセスサービスによってアクセスされるファイル等のデータを管理する少なくとも1つのファイルシステム4と、前述のデータを格納する記憶装置5とをそれぞれ備えており、各クライアント1にファイルアクセスサービスを提供するファイルサーバである。

【0040】図1では、クライアント1とサーバ3とスイッチ100とが、同一のネットワーク2に接続されているが、論理的には、スイッチ100は各クライアント1と各サーバ3との間に配置されている。このようなリモートファイルシステムにおいては、全てのファイルアクセス操作は、クライアント1から送信されるファイルアクセス要求（例えば、いわゆるUnix^(R)系のシステムであればNFSプロトコルによる要求）と、そのファイルアクセス要求に対してサーバ3から返信される応答とで成り立っており、サーバ3から何らかの要求が出力されることはないものとする。

【0041】クライアント1から送信されるファイルアクセス要求には、アクセス対象となるディレクトリやファイルなどのオブジェクトを識別するためのオブジェクトIDが含まれている。クライアント1は、ディレクトリやファイルなどのオブジェクトを識別するためのオブジェクトIDをファイルアクセス要求の中に指定することによって、サーバ3によって管理されるオブジェクトへアクセスする。オブジェクトIDはサーバ3によって生成されるものであり、クライアント1は最初にアクセス可能なオブジェクトの一覧を有している。クライアン

ト1は、まず、その一覧から、オブジェクトまでのアクセスパスを指定した要求データを送信し、最初にアクセス可能なオブジェクトのオブジェクトIDをサーバ3の応答データから取得する。また、オブジェクトIDをすでに取得しているオブジェクトの下位に配置されたオブジェクトについては、クライアント1は、すでに取得しているオブジェクトIDと、ファイルアクセスを行うオブジェクトの名前を指定した要求データを送信し、その応答データからオブジェクトIDを取得する。つまり、クライアント1は、サーバ3からの応答データを受信して、その応答データに含まれるオブジェクトIDを取得しない限り、そのオブジェクトIDに対応するオブジェクトを操作することができない。

【0042】このネットワークファイルシステムでは、サーバ3の存在は、スイッチ100によってクライアント1から隠蔽されており、クライアント1からのファイルアクセス要求はスイッチ100に送信される。スイッチ100は、クライアント1から送信されたファイルアクセス要求を受信し、そのファイルアクセス要求を適切なサーバ3に振り分けて転送（ルーティング）し、サーバ3から送信された前述のファイルアクセス要求に対する応答を受信し、ファイルアクセス要求を送信したクライアント1へとその応答を転送（ルーティング）するように動作する。

【0043】スイッチ100は、クライアント1からファイルアクセス要求を受信したときには、そのファイルアクセス要求に含まれるオブジェクトIDを解析し、その解析結果に基づいて複数のサーバ3の中から、そのファイルアクセス要求を転送すべきサーバ3を決定し、そのファイルアクセス要求をそのサーバ3に転送する。しかし、サーバ3で生成されるオブジェクトIDは、一般的に、生成したサーバ3のみが解釈可能なデータ列によって構成されているため、スイッチ100およびクライアント1では、そのデータを解釈することができない。また、異なるサーバ3でまったく同一のデータ列を有するオブジェクトIDを生成する可能性もある。そのため、このままでは、スイッチ100がファイルアクセス要求の適切なルーティングを実行することができない。また、前述のように、クライアント1は、サーバ3からの応答データを受信して、その応答データに含まれるオブジェクトIDを取得しない限り、そのオブジェクトIDに対応するオブジェクトを操作することができない。

【0044】そこで、本実施形態のスイッチ装置であるスイッチ100では、サーバ3から送信される応答データ中に含まれるオブジェクトIDに送信元のサーバ3を他のサーバ3から識別するためのサーバ識別情報を挿入し、サーバ識別情報が挿入されたオブジェクトIDを含む応答データによってパケットを再構成し、そのパケットをクライアント1に転送する。このようにすれば、クライアント1は、サーバ識別情報が挿入されたオブジェ

クトIDを含むファイルアクセス要求をスイッチ100に送信するようになり、スイッチ100では、そのオブジェクトIDに挿入されたサーバ識別情報を参照してそのファイルアクセス要求を転送すべきサーバを特定することができる。なお、スイッチ100は、ファイルアクセス要求をサーバ3に転送する際には、サーバ識別情報を挿入したオブジェクトIDを、サーバ3が生成した元のオブジェクトIDに変換する必要がある。このようにスイッチ100を備えることによって、ネットワークファイルシステムにおける複数のサーバによる分散処理が可能となる。

【0045】図2は、スイッチ100の構成を示すブロック図である。図2に示すように、スイッチ100は、オブジェクトID書き換え部101と、ファイルアクセス管理部102と、パケット処理部103とを備えている。パケット処理部103は、ネットワーク2を介してクライアント1およびサーバ3から受信したパケットを解析して、パケットに含まれるデータをファイルアクセス管理部120に出力し、ファイルアクセス管理部120から出力されたデータをパケット化して、ネットワーク2を介してクライアント1およびサーバ3に送信する。ファイルアクセス管理部102は、パケット処理部103から出力されたデータを入力して解析し、パケット処理部103が受信したパケットの転送先を決定する。オブジェクトID書き換え部101は、ファイルアクセス管理部102が決定するパケットの転送先に応じて、前述のデータに含まれるオブジェクトIDの書き換えを行う。

【0046】図3は、ファイルアクセス管理部102の構成を示すブロック図である。図3に示すように、ファイルアクセス管理部102は、転送手段（第2の転送手段を含む）の一例であるデータ解析部124と、第2のテーブルの一例であるトランザクションID管理テーブル125と、サーバ情報管理テーブル126とを備えている。データ解析部124は、要求データおよび応答データを解析して、受信したパケットの適切な送信先を決定する。トランザクションID管理テーブル125には、各要求および応答データを関連づけ、個々の要求および応答データをクライアント1およびサーバ3で区別するために付与されているトランザクションIDが登録されている。サーバ管理テーブル126には、サーバ識別情報と、サーバ3のIPアドレスやMACアドレス等のサーバのアドレス情報との組み合わせがエントリされる。

【0047】図4は、オブジェクトID書き換え部101の構成を示すブロック図である。図4に示すように、オブジェクトID書き換え部101は、要求応答分配部120と、要求処理部121と、応答処理部122と、第3のテーブルの一例であるアルゴリズムテーブル123とを備えている。要求応答分配部120は、ファイル

10

20

30

40

50

アクセス管理部 102 から送信されたデータが要求データである場合には、そのデータを要求処理部 121 に分配し、応答データである場合には、そのデータを応答処理部 122 に分配する。アルゴリズムテーブル 123 には、データの書き換え方が登録されている。要求処理部 121 および応答処理部 123 は、アルゴリズムテーブル 123 に登録された書き換え方に基づいて、オブジェクト ID の書き換えを行い、書き換えたオブジェクト ID を要求応答分配部 120 に送信する。要求応答分配部 120 は、書き換えられたオブジェクト ID をファイルアクセス管理部 102 に返信する。

【0048】＜オブジェクト ID フォーマット＞図 5 は、オブジェクト書き換え部 101 において書き換えられるオブジェクト ID のフォーマットの一例を示す図である。図 5 (a) は、サーバ 3 によって生成されるオブジェクト ID、すなわち元オブジェクト ID 150 である。元オブジェクト ID 150 は、例えば、NFS 等のネットワークファイルアクセスプロトコルの種類やバージョンによってデータ長の定義が異なる。つまり、元オブジェクト ID 150 は、一般的に、データ長が固定であるものと、最大データ長のみが決まっており、最大データ長の範囲内でデータ長が可変であるものの 2 種類に分類される。図 5 (b) は、データ長が固定であるオブジェクト ID であって、オブジェクト書き換え部 101 によってサーバ識別情報 154 およびアルゴリズム識別情報 153 が挿入された情報付帯オブジェクト ID のフォーマットが示されている。図 5 (b) に示すように、このオブジェクト ID はデータ長が固定であるので、サーバ識別情報 154 およびアルゴリズム識別情報 153 が挿入されても、全体のデータ長は固定（元オブジェクト ID 150 と同じ長さ）となっていなければならない。したがって、オブジェクト ID 書き換え部 101 は、図 5 (b) に示す情報付帯オブジェクト ID を生成する際には、図 5 (a) に示す元オブジェクト ID 150 の一部を削除して圧縮データ 152 とし、サーバ識別情報 154 およびアルゴリズム識別情報 153 を先頭に付与する。なお、アルゴリズム識別情報 153 とは、元オブジェクト ID 150 の一部を削除する際にオブジェクト ID 書き換え部 101 によって用いられる削除アルゴリズムを識別するための情報である。このように書き換えられた情報付帯オブジェクト ID を、圧縮型情報付帯オブジェクト ID 151 と呼ぶ。

【0049】図 5 (c) は、データ長が可変長であるオブジェクト ID であって、サーバ識別情報およびアルゴリズム識別情報が挿入された情報付帯オブジェクト ID が示されている。図 5 (c) では、元オブジェクト ID 150 のデータは削除されずに、サーバ識別情報 154 およびアルゴリズム識別情報 153 がそのまま追加されている。このような情報付帯オブジェクト ID を追記型情報付帯オブジェクト ID 155 と呼ぶ。なお、データ

長が可変長であっても、最大のデータ長を越える場合には、オブジェクト ID 書き換え部 101 では、追記型情報付帯オブジェクト ID 155 でなく、圧縮型情報付帯オブジェクト ID 152 が生成される。

【0050】また、図 5 (b) に示すサーバ識別情報 154 およびアルゴリズム識別情報 153 とは、図 5

(c) に示すサーバ識別情報 154 およびアルゴリズム識別情報 153 と、それぞれデータ長が同じであるのが望ましく、サーバ識別情報 154 とアルゴリズム識別情報 153 とは、それぞれ同じデータ位置に設定されるのが望ましい。オブジェクト ID の大きさは、一般的に NFS バージョン 2 プロトコルでは、32 バイト程度であり、サイズが大きい。また、NFS バージョン 3 では、そのサイズを変えることもある。そのため、その中からデータ長や位置がまちまちである情報を抽出および検索するには時間を要する。したがって、情報付帯オブジェクト ID に含まれる各識別情報の抽出や検索するには、それらのデータを同じデータ長、同じ位置とするのが望ましい。

【0051】図 6 は、アルゴリズムテーブル 123 の構成を示す図である。アルゴリズムテーブル 123 は、アルゴリズム識別情報 153 と、復元情報の一例である削除データ 156 と、同じく復元情報の一例である削除位置 157 との組によって構成されたエントリを有するテーブルである。このテーブルを参照すれば、要求処理部 121 は、削除データ 156 と、削除データ位置 157 とにより、圧縮データ 152 を元オブジェクト 150 へ戻すことが可能になる。なお、アルゴリズムテーブル 123 の削除データ 156 と、削除データ位置 157 の情報のないエントリは、圧縮データ 152 を生成せず、元オブジェクト 150 にそのままサーバ識別情報 154 およびアルゴリズム識別情報 153 を追加した追記型情報付帯オブジェクト ID のエントリである。

【0052】なお、サーバ 3 において元オブジェクト ID 150 が一定の生成規則で生成されている場合には、元オブジェクト ID 150 は、データ列に周期性のある（各元オブジェクト ID 150 中に同一のビット列が同一の場所に発生する）ID となる。このような場合には、削除データ 156 と、削除データ位置 157 とが等しくし、同じアルゴリズム識別情報 153 を用いることによって、アルゴリズムテーブル 123 のエントリ数を削減することも可能である。

【0053】＜要求パケット受信の際の動作＞次に、スイッチ 100 における、クライアント 1 から送られる要求パケットを受信した後、複数のサーバ 3 の中から適切なサーバを選択し、そのサーバにパケットを送信する場合の動作の手順について説明する。

【0054】パケット処理部 103 は、クライアント 1 から送信された要求パケットを受信すると、そのパケットからファイルアクセス要求データ、クライアント 1 の

クライアントIPアドレス情報160およびクライアントMACアドレス情報161を抽出し、それらのデータをファイルアクセス管理部102に送信する。

【0055】ファイルアクセス管理部102のデータ解析部124は、それらのデータを受信すると、ファイルアクセス要求データを解析して、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155を抽出する。

【0056】そして、データ解析部124は、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155に含まれるサーバ識別情報154を抽出して、図7に示すサーバ情報管理テーブル126を参照し、サーバ識別情報154に対応するサーバ3のサーバIPアドレス情報158とサーバMACアドレス情報159とを読み出す。

【0057】ファイルアクセス要求データには、各要求データを区別するためのトランザクションID160が含まれている。データ解析部124は、トランザクションID160と、バケット管理部103より受信したクライアント1のクライアントIPアドレス情報161およびクライアントMACアドレス情報162と、ファイルアクセス要求データに含まれる、リードやライトなどのファイルアクセス命令情報163と、サーバ識別情報154とを、図8に示すトランザクションID管理テーブル125に登録する。

【0058】その後、データ解析部124は、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155を、情報付帯オブジェクトID書き換え命令とともに、オブジェクトID書き換え部101に送信する。これらのオブジェクトIDは、オブジェクトID書き換え部101において元オブジェクトID150に変換された後、再びファイルアクセス管理部102に送信される。

【0059】ファイルアクセス管理部102は、元オブジェクトID150をファイルアクセス要求に組み込む。ファイルアクセス管理部102において、再構成されたファイルアクセス要求データと、サーバ情報管理テーブル126を検索することにより特定された、サーバ3のサーバIPアドレス情報158、およびサーバMACアドレス情報159とは、バケット処理部103に送信される。バケット処理部103は、それらを要求バケットとして再構成し、そのバケットをサーバ3に送信する。

【0060】＜要求データの受信の際のオブジェクトID書き換え部の動作＞次に、クライアント1からの要求データに含まれる、圧縮型情報付帯オブジェクトIDもしくは追記型情報付帯オブジェクトID155を、元オブジェクトID150に復元する際の、オブジェクトID書き換え部101における動作の手順について説明する。

【0061】要求データの解析後に、ファイルアクセス処理部102から送信された、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155と、情報付帯オブジェクト書き換え命令を受信すると、要求応答分配部120は、情報付帯オブジェクトID書き換え命令に基づいて、受信した圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155を要求処理部121に送信する。

10 【0062】要求処理部121は、アルゴリズムテーブル123を参照して、受信した情報付帯オブジェクトIDに含まれるアルゴリズム識別情報153に対応するエントリの削除データ156および削除データ位置157を特定し、それらに基づいて元オブジェクトID150を生成し、要求応答分配部120を介して、ファイルアクセス管理部102に生成した元オブジェクトID150を送信する。

20 【0063】＜応答バケット受信の際の動作＞次に、サーバ3から送られる応答バケットを、複数のクライアント1の中から適切なクライアントを選択して転送する際のスイッチ100における動作の手順について説明する。バケット処理部103は、受信したサーバ3からの応答バケットから、応答データ、サーバ3のサーバIPアドレス情報158もしくはサーバMACアドレス情報159等のサーバアドレス情報を抽出し、ファイルアクセス管理部102に送信する。

30 【0064】ファイルアクセス管理部102のデータ解析部124は、サーバ管理テーブル126を参照し、サーバIPアドレス情報158もしくはサーバMACアドレス情報159と等しい値を持つエントリを検索し、そのエントリのサーバ識別情報154を得る。さらに、データ解析部124は、応答データを解析することにより、応答データに含まれるトランザクションID160を抽出し、トランザクションID管理テーブル125を参照して、サーバ識別情報154とトランザクションID160がともに一致するエントリを検索し、受信した応答データに対応するファイルアクセス要求を送信したクライアント1のクライアントIPアドレス情報161と、クライアントMACアドレス情報162と、ファイルアクセス命令情報163とを得る。

40 【0065】サーバ3からの応答バケットには、オブジェクトIDが含まれていない場合もある。特定のファイルアクセスサービスのプロトコルやバージョンでは、ファイルアクセス命令情報が応答データに明示されていない場合があるため、データ解析部124は、ファイルアクセス命令情報163を用いて、どの種類の命令に対する応答データかを特定することにより、応答データ中に元オブジェクトIDが含まれているかどうかを判断することができる。

50 【0066】応答データに元オブジェクトID150が

含まれている場合は、データ解析部 124 は、元オブジェクト ID 150 とともに、データ長の定義方法と、サーバ識別情報と、元オブジェクト書き換え命令とをオブジェクト ID 書き換え部 101 に送信する。オブジェクト ID 書き換え部 101 は、受信した元オブジェクト ID を、圧縮型情報付帯オブジェクト ID 151 もしくは追記型情報付帯オブジェクト ID 155 に書き換え、データ解析部 124 は、応答データを再構成する。データ解析部 124 は、応答データに元オブジェクト ID が含まれていない場合には、応答データに変更を加えない。その後、データ解析部 124 は、クライアント IP アドレス情報 161 と、クライアント MAC アドレス情報 162 と、応答データをバケット処理部 103 に送信する。バケット処理部 103 は、応答バケットを生成し、応答バケットをクライアント 1 に送信する。

【0067】＜応答データ受信の際のオブジェクト書き換え部の動作＞図 9 は、応答データを受信した際のオブジェクト書き換え部 101 の動作を示すフローチャートである。図 9 に示すように、要求応答分配部 120 は、
10 応答データに含まれる、元オブジェクト ID 150 とともに、データ長の定義方法と、サーバ識別情報と、元オブジェクト書き換え命令をファイルアクセス管理部 120 から受信すると、元オブジェクト書き換え命令によって、応答処理部 122 へ受信したデータをすべて分配する (S100)。

【0068】応答処理部 122 は、受信したデータの定義方法を判別し、元オブジェクト ID のデータ長が固定長であるか否かチェックする (S101)。ステップ S101 において、固定長である場合には、応答処理部 122 は、元オブジェクト ID のデータ列の一部を削除して圧縮データを生成し (S102)、削除データ 156
20 と、削除データ位置 157 と、削除アルゴリズムを区別するために付けられたアルゴリズム識別情報 153 を、アルゴリズムテーブル 123 に登録する (S103)。

【0069】アルゴリズムテーブル 123 への登録終了後、応答処理部 122 は、サーバ識別情報 154 と、アルゴリズム識別情報 153 と、圧縮データ 152 を組み合わせて、圧縮型情報付帯オブジェクト ID 151 を生成し (S104)、その圧縮型情報付帯オブジェクト ID 151 を、要求応答分配部 120 を経由して、
30 ファイルアクセス管理部 102 へ送信する (S105)。

【0070】ステップ S101 において、元オブジェクト ID 150 のデータ長が固定長でない場合には、応答処理部 122 は、サーバ識別情報 154 と、アルゴリズム識別情報 153 と、元オブジェクト ID 150 の 3 つのデータ長の和を計算し、計算値と定義されている最大データ長を比較する (S106)。

【0071】ステップ S106 において、最大データ長よりも 3 つのデータ長の和が長い場合には、応答処理部 122 は、データ長が固定長である場合と同様の手順で
50

処理を行い (S102~S105)、3 つのデータ長の和が最大データ長よりも短い場合は、削除データ 156 および削除データ位置 157 の情報が登録されないエンタリとしてのアルゴリズム識別情報 153 を生成し、アルゴリズムテーブル 123 へ登録する (S107)。

【0072】応答処理部 122 は、アルゴリズムテーブル 123 への登録終了後、サーバ識別情報 154 と、アルゴリズム識別情報 153 と、元オブジェクト ID 150 を組み合わせて、追記型情報付帯オブジェクト ID 155 を生成し (S108)、要求応答分配部 120 を経由して、追記型情報付帯オブジェクト ID 155 をファイルアクセス管理部 102 に送信する (S105)。

【0073】なお、本実施形態のスイッチ装置では、サーバ 3 で公開 (エクスポート) する複数のオブジェクトの公開単位、すなわち共有オブジェクト単位で分割されている場合には、そのオブジェクトがその共有オブジェクト単位に含まれるかを識別するための共有オブジェクト単位 ID の項目をサーバ情報管理テーブル 126 のエンタリとして追加してもよい。このようにすれば、スイッチ 100 がオブジェクトを格納するサーバ 3 だけでなく、オブジェクトが含まれる共有オブジェクト単位も把握することができるようになるため、スイッチ 100 においてサーバ単位よりもさらに分解能が高い共有オブジェクト単位でのきめ細かい転送管理を実現することが可能となる。

【0074】(第 2 の実施形態) 次に、本発明の第 2 の実施形態のスイッチ装置について説明する。本実施形態のスイッチ装置は、第 1 の実施形態のスイッチ装置と同様に、図 1 に示すネットワークファイルシステムに用いられる装置である。

【0075】図 10 は、本実施形態のスイッチ装置の構成を示すブロック図である。図 10 に示すように、本実施形態のスイッチ 100 は、第 1 の実施形態のスイッチ装置において備えられているオブジェクト ID 書き換え部 101 と、ファイルアクセス管理部 102 と、バケット処理部 103 とに加え、擬似ファイルシステム 105 を備えている。擬似ファイルシステム 105 は、複数のサーバ 3 のファイルシステム 4 で構成されている複数のディレクトリツリーを組み合わせ、1 つの統合ディレクトリツリーとして管理するためのシステムである。ファイルシステム補完処理部 104 は、クライアント 1 から
40 のファイルアクセス要求が、複数サーバ 3 もしくは複数のファイルシステム 4 等の複数の送信先に渡ってファイルアクセスを必要とする場合の補完処理を行う。

【0076】図 11 は、本実施形態のスイッチ装置におけるファイルアクセス管理部 102 の構成を示すブロック図である。図 11 に示すように、本実施形態のスイッチ装置におけるファイルアクセス管理部 102 は、第 1 の実施形態のスイッチ装置において備えられているデータ解析部 124 と、トランザクション ID 管理テーブル

125と、サーバ情報管理テーブル126とに加え、第4のテーブルの一例である特殊オブジェクトID管理テーブル127を備えている。特殊オブジェクトID管理テーブル127は、複数のサーバ3もしくは複数のファイルシステム4等の複数の送信先に、ファイルアクセスが必要となるような、クライアント1からのファイルアクセス要求を判断するためのテーブルである。さらに、本実施形態のスイッチ装置におけるデータ解析部124は、オブジェクトID書き換え部101やパケット処理部103とのデータ入出力だけでなく、ファイルシステム補完処理部104とのデータの入出力が可能となっている。

【0077】次に、本実施形態のスイッチ装置であるスイッチ100の動作について説明する。本実施形態のスイッチ装置は、第1の実施形態のスイッチ装置における要求パケットおよび応答パケットを、クライアント1もしくはサーバ3へ適切に振り分ける機能に加え、複数のサーバ3のファイルシステム4上に形成されている複数のディレクトリツリーを擬似ファイルシステム105のディレクトリツリーに統合することによって、複数のサーバ3における個々のファイルシステム4の存在をクライアント1に意識させることなく、スイッチ100の擬似ファイルシステム105のみにアクセスしているかのように振舞わせる機能を有している。

【0078】擬似ファイルシステム105は、サーバ3によりネットワーク2を介してアクセス可能なものとして公開されているファイルシステム4上の複数のディレクトリツリーにおけるツリー構造を、擬似ファイルシステム105上で自由に組み合わせ、1つのディレクトリツリーにマッピングする。

【0079】スイッチ100は、クライアント1から受信したファイルアクセス要求パケットを、単に特定のサーバ3へ振り分けるだけでなく、複数のサーバ3もしくは異なるファイルシステム4にまたがるようなファイルアクセス要求を受け取った場合に、1つのファイルアクセス要求を、複数のファイルアクセス要求に分割して複数のサーバ3もしくはファイルシステム4へ送信し、複数のサーバ3もしくはファイルシステム4から送信された全ての応答データをスイッチ100で1つの応答データに再構成してクライアント1に返信する。

【0080】図12は、擬似ファイルシステム105のディレクトリツリー200の一例を示す図である。図12に示すように、擬似ファイルシステム105は、サーバA(250)で公開されているディレクトリb(221)、およびサーバB(251)で公開されているディレクトリc(221)からのディレクトリツリーのツリー構造を、擬似ファイルシステム105上に形成されたルートディレクトリ220のサブディレクトリからのディレクトリツリーとしてマッピングする。さらに、擬似ファイルシステム105は、サーバB(251)で公開

されているディレクトリe(224)からのディレクトリツリーのツリー構造を、ディレクトリb(221)からのディレクトリツリーの下層に存在するディレクトリf(223)のサブディレクトリからのディレクトリツリーとして登録することによって、擬似ファイルシステムディレクトリツリー200を生成する。

【0081】なお、本実施形態のスイッチ装置では、擬似ファイルシステムディレクトリツリー200上におけるディレクトリ名は、サーバ3のファイルシステム4で設定されたディレクトリ名と必ずしも同一にする必要はなく、別のディレクトリ名を設定して、クライアント1に公開してもよい。

【0082】擬似ファイルシステム105は、擬似ファイルシステム105のディレクトリツリー200のツリー構造のみを管理しており、オブジェクトのデータ、属性情報、各ファイルシステム4の接合部以外のツリー構造等は、すべてサーバ3が管理を行っている。したがって、スイッチ100では、複数のファイルシステム4を結合しているツリー部分のみの名前解決処理を行い、その他のツリー部分については、サーバ3に処理を任せている。このようにすれば、名前解決処理の負荷を、スイッチ100とサーバ3との間で分散することができる。

【0083】スイッチ100は、例えば、クライアント1から、ディレクトリe(224)に対応するオブジェクトIDや属性情報を取得する命令となる、ディレクトリf(223)に対応するオブジェクトIDを用いたファイルアクセス要求パケットを受信した場合には、その命令に対して正しく応答できるように新しい要求パケットを生成し、ディレクトリf(223)のデータを格納しているサーバA(250)と、ディレクトリe(224)のデータを格納しているサーバB(251)へ要求パケットをそれぞれ送信する。そして、その結果サーバA(250)、サーバB(251)から返信されてくる2つの応答パケットを再構成する処理を行うための管理情報として、擬似ファイルシステム105が用いられる。

【0084】スイッチ100は、クライアント1からのファイルアクセス要求により、複数のサーバ3もしくはファイルシステム4へそれぞれ要求データを送るような補完処理を行わなければならない場合以外、第1の実施形態のスイッチと同様の手順で処理を行うが、どちらの処理を行うかは、ファイルアクセス管理部102の特殊オブジェクトID管理テーブル127を用いて判断する。

【0085】特殊オブジェクトID管理テーブル127は、図13に示すように、ディレクトリb(221)や、ディレクトリe(224)などの、ディレクトリツリー200の異なるサーバ3上に存在するツリーのつなぎ目にあたるすべてのディレクトリのオブジェクトID164と、補完処理を行わなければならないすべてのフ

ファイルアクセス命令情報165からなるエントリによって構成されている。特殊オブジェクトID管理テーブル127には、擬似ファイルシステム105のディレクトリツリー200が生成された時点で、対応するすべてのディレクトリが登録されている。なお、特殊オブジェクトID管理テーブル127に登録するオブジェクトID164は、クライアント1から送信される要求データに含まれる情報付帯オブジェクトIDと照合するため、情報付帯オブジェクトID、すなわち圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155である。このようにすれば、照合時間を短縮することができる。

【0086】次に、クライアント1からの要求パケットやサーバ3からの応答パケットを受信した場合のスイッチ100における動作の手順について詳細に説明する。図14は、本実施形態のスイッチ装置の動作を示すフローチャートである。

【0087】図14に示すように、クライアント1からのファイルアクセス要求パケットを受信すると、パケット処理部103は、ファイルアクセス要求データと、クライアント1のクライアントIPアドレス情報160およびクライアントMACアドレス情報161をパケットより抽出し、ファイルアクセス管理部102に送信する(S110)。

【0088】ファイルアクセス管理部102のデータ解析部124は、受信した要求データに含まれるリードやライトといった命令情報と、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155を抽出し、特殊オブジェクトID管理テーブル127に登録されたオブジェクトID164と命令情報165とのエントリにそれらが該当するものがあるか否かを確認する(S111)。

【0089】要求データから抽出した命令情報と、情報付帯オブジェクトID151もしくは155との2つのデータのうち、特殊オブジェクトID管理テーブル127のエントリに、どちらか1つでも該当しない場合、以降のスイッチ100における要求パケットに関する処理、および前記要求パケットに対応する応答パケットに関する処理は、第1の実施形態のスイッチ装置における手順で処理が実行される(S112)。

【0090】要求データから取り出した命令情報と、オブジェクトID151もしくは155の2つのデータのうち、特殊オブジェクトID管理テーブル127のエントリにどちらも該当する場合、データ解析部124は、トランザクションID管理テーブル125に必要となるデータをすべて登録した後、受信した要求データをすべてファイルシステム補完処理部104へ送信する(S113)。

【0091】ファイルシステム補完処理部104は、擬似ファイルシステム105を参照して受信した要求デー

タの解析を行った後、複数のサーバ3もしくはファイルシステム4へ送信する複数の要求データを生成し、パケット管理部103を介して、すべての要求パケットを、関連する複数あるいは1つのサーバ3へ送信する(S114)。なお、本ステップにおいて、1つのサーバ3へ全ての要求パケットの送信するのは、サーバ3が複数のファイルシステム4を備えており、擬似ファイルシステム105上のディレクトリツリー200構成のマッピングが、それぞれのファイルシステム4を接合することで構成されている場合に限られる。

【0092】スイッチ100のパケット処理部103は、サーバ3に送信した要求パケットに対する応答パケットを受信すると、その応答パケットから、応答データやサーバ3のアドレス情報158、159を抽出し、それらをファイルシステム補完処理部104に送信する(S115)。

【0093】ファイルシステム補完処理部104は、応答データに元オブジェクトID150が含まれている場合には、オブジェクトIDの書き換え命令と、元オブジェクトID150と、サーバ3のアドレス情報158、159とをファイルアクセス管理部102へ送信する。

【0094】ファイルアクセス管理部102は、受信したすべての応答データに含まれる元オブジェクトID150をオブジェクトID書き換え部101に送信し、第1の実施形態と同様に、圧縮型情報付帯オブジェクトID151あるいは追記型情報付帯オブジェクトID155に書き換えさせ、その情報付帯オブジェクトIDと、受信したすべての応答データを用いて、クライアント1からの要求データに対する応答データを生成する(S116)。

【0095】ファイルアクセス管理部102は、トランザクションID管理テーブル125に登録されているデータと一致するエントリを検索し、クライアント1のアドレス情報を得た後、応答データとクライアント1のアドレス情報をパケット処理部103に送信する(S117)。パケット処理部103は、受信した応答データとクライアント1のアドレス情報とに基づいて応答パケットを生成し、その応答パケットをクライアント1に送信する。

【0096】(第3の実施形態)次に、本発明の第3の実施形態のスイッチ装置について説明する。本実施形態のスイッチ装置は、第1の実施形態および第2の実施形態と同様に、図1に示すネットワークファイルシステムに適用される。

【0097】図15は、本実施形態のスイッチ装置の構成を示すブロック図である。図15に示すように、本実施形態のスイッチ100は、第1、第2の実施形態のスイッチ装置に備えられているオブジェクトID書き換え部101と、ファイルアクセス管理部102と、パケット処理部103と、第2の実施形態に備えられているフ

10

20

30

40

50

ファイルシステム補完処理部104と、擬似ファイルシステム105とに加え、データ移動処理部106を備えている。データ移動処理部106は、各サーバ3のファイルアクセス処理の負荷や、記憶装置5の残り容量に基づいて、擬似ファイルシステム105上の統合ディレクトリツリーのディレクトリ構造を変更することなく、複数サーバ3間でデータの移動を行う。

【0098】図16は、本実施形態のスイッチ装置におけるファイルアクセス管理部102の構成を示すブロック図である。図16に示すように、ファイルアクセス管理部102は、第1、第2の実施形態のスイッチ装置において備えられているトランザクションID管理テーブル125およびサーバ情報管理テーブル126と、第2の実施形態において備えられている特殊オブジェクトID管理テーブル127に加え、第5のテーブルの一例である移動オブジェクトID管理テーブル128を備えている。移動オブジェクトID管理テーブル128には、複数のサーバ3間でデータを移動した場合に生じるオブジェクトIDの変更による不具合を生じさせないようにする情報がエントリされている。また、移動隠蔽手段の一例を含むデータ解析部124では、オブジェクトID書き換え部101と、パケット処理部103と、ファイルシステム補完処理部104とのデータの入出力だけでなく、データ移動処理部106とのデータの入出力が可能となっている。

【0099】次に、本実施形態のスイッチ装置の動作について説明する。本実施形態のスイッチ装置は、第1、第2の実施形態のスイッチ装置において備えられていた機能に加え、擬似ファイルシステム105上に構成されている統合ディレクトリツリーのツリー構造を保持したまま、その統合ディレクトリツリーを構成している各サーバ3の公開ディレクトリを元とするツリー構造を、別のサーバ3へコピーしてデータ移動を行った後、ツリー構造のマッピングを変更することによって、クライアント1にデータの移動を隠蔽し、統合ディレクトリツリーの構成を不変としたまま、サーバ3の記憶装置5における容量不足や、各サーバ3におけるファイルアクセス処理の負荷を軽減させる。

【0100】サーバ3間におけるデータ移動に伴い、各ディレクトリやファイルなどのオブジェクトに対応するオブジェクトIDも変更する必要があるため、スイッチ100では、データ移動前とデータ移動後のオブジェクトIDの変更履歴を管理し、クライアント1およびサーバ3において、オブジェクトIDの変更による不具合が生じないようにすることによって、データの移動の操作をクライアント1から隠蔽する。

【0101】例えば、図12に示されている擬似ファイルシステム105では、サーバB(251)における記憶装置の残り容量が少なく、サーバA(250)の記憶装置の残り容量に余裕があり、サーバB(251)のデ

ィレクトリc(222)からのディレクトリツリーにおけるデータと、オブジェクトの所有者、アクセス権などの属性情報を、ツリー構造ごとサーバA(250)にコピーするのが望ましい状況であるとする。本実施形態のスイッチ装置は、図17に示すように、ディレクトリc(225)からのディレクトリツリーを擬似ファイルシステム105のディレクトリツリー200の構造を変えことなくマッピングし直すように擬似ファイルシステム105の情報を更新した後、サーバB(251)のディレクトリc(222)からのディレクトリツリーのデータを削除する。このようにすれば、本実施形態のスイッチ装置では、クライアント1にデータの移動を隠蔽しつつ残り容量を増やすことができるようになるため、ディレクトリc(225)からのディレクトリツリーやディレクトリe(224)からのディレクトリツリーにおけるデータの追加が可能となる。

【0102】次に、データ移動操作に関するスイッチ100での動作手順について説明する。データ移動処理部106は、すべてのサーバ3における記憶装置5の残り容量データの監視と、スイッチ100で処理されているファイルアクセス要求パケットや応答パケットに関する履歴データを記録している。そして、データ移動管理部106は、あらかじめ設定しておいた基準よりも、残り容量が少なくなった場合や、特定のサーバ3のみにファイルアクセス要求が集中していた場合に、擬似ファイルシステム105に登録されているデータを参照し、移動元のディレクトリツリーと、移動先を決定し、データのコピーを開始する。

【0103】データコピーを実行する際に、データ移動処理部106は、クライアント1が既に取得している情報付帯オブジェクトID151、155により、移動データの対象となっているデータにアクセスする場合の不具合に対応するために、データの移動の対象となっているすべてのオブジェクトに関連する情報として、移動元のサーバ3が生成した元オブジェクトID150をオブジェクトID書き換え部101で書き換えた情報付帯オブジェクトID151、155と、移動先のサーバ3で生成された元オブジェクトID150と、サーバ情報管理テーブル126に登録されている移動先のサーバ3に対応したサーバ識別情報154とを、図18に示す移動オブジェクトID管理テーブル128へ登録する。図18に示すように、移動オブジェクトID管理テーブル128は、移動元オブジェクトID166と、移動先オブジェクトID167と、移動先サーバ情報168の情報からなるエントリにより構成されている。

【0104】移動オブジェクトID管理テーブル128に登録されているオブジェクトが再び移動する場合には、移動オブジェクトID管理テーブル128の移動先オブジェクトID167が移動元オブジェクトID166の項目に追加され、残りの移動先オブジェクトID1

10

20

30

40

50

67と移動先サーバ情報168の項目は最新の状態に更新される。

【0105】移動先オブジェクトID167に登録された情報のうち、ファイルアクセス補完処理部104での処理が必要となる可能性のあるオブジェクトを示すオブジェクトIDとしては、元オブジェクト150ではなく、圧縮型情報付帯オブジェクトID151および追記型情報付帯オブジェクトID155を登録する。この場合には、移動先サーバ情報168へはサーバ識別情報154を登録しないでおく。

【0106】すべてのデータのコピーが完了し、移動元のサーバ3のデータを削除した後、データ移動処理部106は、移動先のサーバ3でコピーを行ったディレクトリツリーに対して、ネットワーク2を介したファイルアクセスを可能とする公開ディレクトリの設定が行われていない場合は、移動先のサーバ3へ公開ディレクトリの設定を行って、擬似ファイルシステム105におけるツリー構造のマッピング情報の更新と、特殊オブジェクトID管理テーブル127に登録されているエントリのデータの更新を行う。

【0107】次に、データ移動が行われた後で、クライアント1からの要求バケットおよびサーバ3からの応答バケットを受信した場合のスイッチ100における詳細な処理手順について詳細に説明する。図19は、バケットを受信した場合のスイッチ100の動作を示すフローチャートである。

【0108】クライアント1からのファイルアクセス要求バケットを受信すると、バケット処理部103は、ファイルアクセス要求データと、クライアント1のクライアントIPアドレス情報160およびクライアントMACアドレス情報161がバケットから抽出し、それらをファイルアクセス管理部102に送信する(S120)。

【0109】データ解析部124は、受信した要求データを解析し、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155を抽出し、そのIDが移動オブジェクトID管理テーブル128に登録されているか否かを確認する(S121)。

【0110】ステップS121において、抽出された圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155が、移動オブジェクトID管理テーブル128の移動元オブジェクトID166として登録されていない場合には、以降の要求データ、およびその要求データに対応する応答データに関する処理は、第2の実施形態のステップS111からの処理が実行される(S122)。

【0111】ステップS121において、抽出された圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155が、移動オブジェクトID管理テーブル128の移動元オブジェクトID166

として登録されている場合には、データ解析部124は、移動先サーバ情報168にサーバ識別情報154が登録されているかを確認する(S123)。

【0112】ステップS123における確認後、移動先サーバ情報168にサーバ識別情報154が登録されていない場合には、要求データに含まれていた圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155を、移動先オブジェクトID167に登録されている圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155に変更し(S124)、以降の要求データ、およびその要求データに対応する応答データに関する処理は第2の実施形態のステップS111からの処理が実行される(S125)。

【0113】ステップS123における確認後、移動先サーバ情報168にサーバ識別情報154が登録されている場合には、データ解析部124は、要求データに含まれていた圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155を、移動先オブジェクトID167として登録されている元オブジェクトID150に変更し、移動先サーバ情報168に登録されているサーバ識別情報154を抽出する(S126)。

【0114】データ解析部124は、トランザクションID管理テーブル125のエントリに必要な情報をすべて登録した後、移動オブジェクトID管理テーブル128から得られたサーバ識別情報154に基づいて、サーバ情報管理テーブル126から、サーバIPアドレス情報158、およびサーバMACアドレス情報159を抽出し、バケット処理部103へ、元オブジェクトID150に変更した要求データと、サーバIPアドレス情報158と、サーバMACアドレス情報159とを送信する(S127)。

【0115】バケット処理部103は、ファイルアクセス管理部102から送信されたデータを用いて要求バケットを生成し、サーバ3に送信する(S128)。

【0116】その要求バケットに対するサーバ3からの応答バケットに関するスイッチ100の処理は、第1の実施形態のスイッチでの応答バケットの処理が実行される(S129)。

【0117】(第4の実施形態) 次に、本発明の第4の実施形態のスイッチ装置について説明する。本実施形態のスイッチ装置は、第3の実施形態のスイッチ装置を变形したものである。第3の実施形態のスイッチ装置では、一度にデータ移動するオブジェクトの数が非常に多い場合に、ファイルアクセス管理部102内の移動オブジェクトID管理テーブル128のエントリ数も非常に多くなって、それらのエントリを検索する時間が増大するようになる。

【0118】そこで、本実施形態のスイッチ装置では、

図 21 に示すように、アルゴリズム識別情報 153 およびサーバ識別情報 154 とともに、データが移動したことを示すデータ移動フラグ 169 をオブジェクト ID に組み込み、オブジェクト ID 書き換え部 101 は、フラグ付き圧縮型情報付帯オブジェクト ID 171、もしくはフラグ付き追記型情報付帯オブジェクト ID 172 を生成する際に、データ移動の対象となったオブジェクトに関しては、データ移動フラグ 169 を真に設定する。

【0119】そして、本実施形態のスイッチ装置では、クライアント 1 からの要求データに含まれるフラグ付き圧縮型情報付帯オブジェクト ID 171 もしくはフラグ付き追記型情報付帯オブジェクト ID 172 を、データ解析部 124 において解析する際に、データ移動フラグが真である場合に限り、データ移動オブジェクト ID 管理テーブル 128 を検索する。このようにすれば、本実施形態のスイッチ装置では、第 3 の実施形態のスイッチ装置よりも効率のよい検索処理を行うことが可能となる。

【0120】（第 5 の実施形態）次に、本発明の第 5 の実施形態のスイッチ装置について説明する。本実施形態のスイッチ装置は、第 3 の実施形態のスイッチ装置を变形したものである。第 3 の実施形態のスイッチ装置では、一度にデータ移動するオブジェクトの数が非常に多い場合に、ファイルアクセス管理部 102 内の移動オブジェクト ID 管理テーブル 128 のエントリ数が非常に多くなって、エントリを検索する時間が増大するようになる。

【0121】そこで、本実施形態のスイッチ装置では、1 回のデータ移動を 1 世代として世代番号をつけ、1 回のデータ移動が終了するたびに世代を 1 つずつ増やし、現在の世代番号をデータ解析部 124 で記憶しておき、移動オブジェクト ID 管理テーブル 128 では、世代番号ごとにテーブルが作成される。

【0122】さらに、オブジェクト ID 書き換え部 101 では、図 22 に示すように、アルゴリズム識別情報 153 およびサーバ識別情報 154 とともに、データ移動世代番号 172 をオブジェクト ID に組み込むようにし、世代番号付き圧縮型情報付帯オブジェクト ID 173 もしくは世代番号付き追記型情報付帯オブジェクト ID 174 を生成する。

【0123】データ解析部 124 は、クライアント 1 からの要求データに含まれる、世代番号付き圧縮型情報付帯オブジェクト ID 173 もしくは世代番号付き追記型情報付帯オブジェクト ID 174 をデータ解析部 124 にて解析して得られた世代番号に基づいて、移動オブジェクト ID 管理テーブル 128 におけるその世代番号のテーブルのみを検索する。このようにすれば、本実施形態のスイッチ装置は、第 3 の実施形態のスイッチ装置よりも効率の良い検索処理を行うことが可能となる。

【0124】（第 6 の実施形態）次に、図 15 および図

23 から図 25 を参照して、本発明の第 6 の実施形態を説明する。ここでは、本発明による擬似ファイルシステムのディレクトリツリーを具体的に説明する。図 23 は、本発明による擬似ファイルシステムのディレクトリツリー 200 を具体的に構成した場合の例を示している。サーバ A (250) のディレクトリ f (223) の下に、サーバ B (251) にあるディレクトリ e (224) を連結する場合、第 6 の実施形態では、ディレクトリ f (223) の下にダミーディレクトリ e' が作成される。そして、擬似ファイルシステム 105 のツリー構成情報より、ファイルアクセス管理部 102 の特殊オブジェクト ID テーブル 127 に、ディレクトリ f (223) の下のオブジェクトを検索するコマンド (NFS の Look up など) がクライアントから発行された場合をピックアップできるようにディレクトリ f (223) のオブジェクト ID が登録される。その結果、当該要求がクライアントより発行された場合、サーバ B (251) のディレクトリ e (224) が擬似ファイルシステムのディレクトリツリー 200 通りに構成されているようにクライアントに見せるため、ファイルシステム補完処理部 104 はサーバ A (250) とサーバ B (251) に要求を分割送付して、その応答より必要な情報を集めて、クライアントに応答を返している。また、属性設定のコマンドがディレクトリ e (224) に発行された場合も特殊処理が必要なため、ディレクトリ e (224) のオブジェクト ID も特殊オブジェクト ID テーブル 127 に登録されている。この場合は、属性設定のコマンドをダミーディレクトリ e' にも反映させるため、やはり、ファイルシステム補完処理部 104 はサーバ A (250) とサーバ B (251) に要求を分割送付する。

【0125】いま、クライアントが、ディレクトリ f (223) を、サーバ A (250) のディレクトリ g の下に移動させる要求を出したとする（この場合は管理者による移動ではないため、実際に擬似ファイルシステムの構成も変化する）。図 24 は、従来のファイルシステムの動作にならって、本発明によるスイッチ装置が移動を実施した結果を示している。この場合、ディレクトリ g の下にディレクトリ f (223) が移動されて、ディレクトリ新 f (300)（実際のディレクトリの名前が変わるわけではないが、対応するオブジェクト ID は変わるため区別の目的で新 f とした）が生成され、その下に見えていたディレクトリ e (224) は、サーバ B (251) からその実態が実際にサーバ A (250) へ移動されて、特殊オブジェクト ID テーブル 127 の内容がクリアされる。これは正常な動作結果であり、クライアントは自分の操作が完遂されたと認識するが、サーバ間でのファイル移動が発生するため、移動させるファイルの分量によっては、完了までに非常に長い時間がかかることがあった。

【0126】これに対し図25は、第6の実施形態による同じ移動の結果を示している。この場合、ディレクトリ f (223) が、ディレクトリ新 f (300) となるところは図24の例と同じであるが、次に、データ移動処理部106のクライアント操作対応部は、このディレクトリ新 f (300) の下に同時にダミーディレクトリ e' (225) を新しいダミーディレクトリ新 e' (301) として移動する。そして、特殊オブジェクトIDテーブル127の内容のうち、登録されているオブジェクトIDをディレクトリ f (223) のものからディレクトリ新 f (300) のものに変更する。これにより、同様にクライアントの要求は完遂されるが、サーバB (251) のディレクトリ e (224) 配下のファイルは移動する必要がないため、非常に高速に移動を完了することができる。

【0127】(第7の実施形態) 次に、本発明の第7の実施形態のスイッチ装置について図面を参照して詳細に説明する。本実施形態のスイッチ装置は、第1の実施形態および第2の実施形態のスイッチ装置を变形したものである。本実施形態は第1の実施形態および第2の実施形態の処理の高速化を目的としている。

【0128】第1の実施形態で、スイッチ100は応答データにオブジェクトIDが含まれている場合にはオブジェクトID書き換え部101でオブジェクトIDを書き換える必要があるため、応答データに含まれるトランザクションIDを元にトランザクションID管理テーブル125からどの種類の命令に対する応答データであるか検索して求めている。また、第2の実施形態では応答パケットがファイルシステム補完処理部104により生成された要求に対する応答であるかを上記と同じくトランザクションID管理テーブル125を検索して求めている。この処理はトランザクションIDがOS Iモデルの第5～7層に属するため、ソフトウェアによりトランザクションIDを抽出してトランザクションID管理テーブル125のエントリと逐次比較をすることになり、時間のかかる処理となっている。

【0129】本実施形態ではハードウェア処理ができるOS Iモデルの第4層以下の情報を用いてどの種類の命令に対する応答であるかを判定できるようにする。例としてOS Iモデルの第4層に属するポート番号を用いた場合を説明する。ここでいうポート番号としては、TCP (transmission control protocol) / IP や UDP (user datagram protocol) でのポート番号が例示される。

【0130】図26は、第7の実施形態でのファイルアクセス管理部を示しており、新たに、分類テーブルの一例である処理分類テーブル129が追加されている。図27に処理分類テーブル129の構成の一例を示す。

【0131】まずサーバ3からの応答パケットに対して行う処理の違いに着目して命令の種類を分けて処理分類情報176を割り当てる。この分け方は命令ごとに分け

てもよいし、処理の違いに着目して同様な処理を行う複数の命令をグループ化してもよい。ただし少なくとも、応答データにオブジェクトIDが含まれない命令グループ、応答データにオブジェクトIDが含まれる命令グループ、ファイルシステム補完処理部で生成された命令グループの3つに分けることにより高速化が可能になる。

【0132】スイッチ100はクライアント1からの要求データをサーバ3に転送する際にどの種類の命令であるかによってスイッチ100からサーバ3へのポート番号を変える。そのポート番号情報175と処理分類情報176を処理分類テーブル129に格納しておく。ポート番号の選択はクライアント1からの要求ごとに空きポート番号を割り当てる方法でもよいし、命令の種類に対してあらかじめ固定的にポート番号を決めておく方法でもよい。ただし前者の方法では命令ごとに処理分類テーブル129のエントリを作ることになるため、処理高速化のためには後者の方法が望ましい。

【0133】サーバ3からの応答パケットはスイッチ100が要求を送ったポート番号に対して送られる。ファイルアクセス管理部102で、サーバ情報管理テーブル126を検索して得たサーバ識別情報154とポート番号情報175をもとに処理分類テーブル129を検索し、処理分類情報176、クライアントIPアドレス情報161、クライアントMACアドレス情報162を得る。ここまでの処理はOS Iモデルの第4層以下をハードウェアで検索できるLSIを使えば高速に処理可能である。

【0134】処理分類情報176から、応答パケットにオブジェクトIDが含まれない命令の場合には上記で得たクライアントIPアドレス情報161、クライアントMACアドレス情報162を元にIPアドレスとMACアドレスを変換するのみで応答パケットを転送することができるためきわめて高速な処理が可能になる。オブジェクトIDが含まれる命令の場合にはオブジェクトID書き換え部101で第1の実施形態と同様な処理を行う。ファイルシステム補完処理部104で生成された命令の場合にはファイルシステム補完処理部104で第2の実施形態と同様な処理を行う。

【0135】このように、処理分類情報176によりサーバ3からの応答パケットを分類し、必要な処理のみ行うようにすることで処理の高速化が可能になる。

【0136】図28で以上のことを実際のハードウェア構成に即したブロックに分けて説明する。データ解析部124は高速処理が可能なデータ解析高速処理部107と処理に時間のかかるデータ解析非高速処理部108に分けることができる。データ解析高速処理部107にはサーバ情報管理テーブル126と処理分類テーブル129がつながり、データ解析非高速処理部108にはトランザクションID管理テーブル125と特殊オブジェクトID管理テーブル127がつながっている。応答デー

タにオブジェクトIDが含まれない命令はポートX、応答データにオブジェクトIDが含まれる命令はポートY、ファイルシステム補完処理部104で生成する命令はポートZを用いてサーバ3に送信する。サーバ3からの応答はポートXの場合はデータ解析高速処理部107だけで処理されるため高速処理が可能となる。ポートYの場合はデータ解析非高速処理部108で処理される。第2の実施形態の場合にファイルシステム補完処理部104を経由していたのに比較して、高速に処理可能である。ポートZの場合はファイルシステム補完処理部104で処理される。この場合もあらかじめファイルシステム補完処理部104で生成された命令に対する応答であることがわかっているため余計な処理が割愛できる効果がある。

【0137】なお本実施形態では、命令の区別にポート番号を用いる場合を例に説明したが、ポート番号以外にもIPアドレス、VLAN(Virtual LAN)タグ、MPLS(multiprotocol label switching)ラベル、MACアドレス等のOSIモデル第4層以下情報であれば同様に実施可能である。

【0138】(第8の実施形態)次に、本発明の第8の実施形態のスイッチ装置について説明する。本実施形態のスイッチ装置は、第2の実施形態のスイッチ装置を変形したものである。本実施形態では第2の実施形態の処理を高速化することを目的としている。

【0139】第2の実施形態ではステップS111においてクライアント1からの要求パケットに含まれるオブジェクトID、命令情報が特殊オブジェクトID管理テーブル127のエントリに該当するかを全ての要求について行う必要がある。オブジェクトIDは例えばNFSでは32バイトの長さがあり、OSIモデルの第5~7層に属するためソフトウェア処理になり時間がかかる。またほとんどの場合には特殊オブジェクトID管理テーブルのエントリに該当しないため、結果的には行わなくてもよい処理を行っていることになる。

【0140】本実施形態ではファイルシステムのつなぎ目となるオブジェクトに特殊な属性を記録しておき、ステップS111の判定を行わずに全ての要求をサーバ3に転送し、応答パケットに含まれる属性が前述の特殊な属性であるかをスイッチ100で判定し、特殊な属性である場合のみファイルシステム補完処理を行うことで処理を高速化する。

【0141】図29を例に説明する。第2の実施形態では、サーバA(250)、サーバB(251)の全てのオブジェクトに対する要求についてオブジェクトIDが特殊オブジェクトID管理テーブル127のエントリに該当するか比較を行う。しかし図29のオブジェクトグループA(226)、オブジェクトグループB(227)はエントリに該当しないため、結果的には行わなくてもよい処理を行っている。そこで本実施形態ではあらか

じめサーバA(250)のディレクトリf(223)の下のダミーのディレクトリe'(225)に特殊な属性を記憶させておく。スイッチ100はクライアント1からのディレクトリe'に対応するオブジェクトIDや属性情報を取得する命令に対して第1の実施形態と同様な処理を行いサーバA(250)に転送する。サーバAはダミーのディレクトリe'(225)のオブジェクトIDと疑似ファイルシステム105により記憶された特殊な属性を含む応答パケットをスイッチ100に送る。スイッチ100は応答パケットに特殊な属性が含まれているかを判定し、この例の場合は含まれているのでファイルシステム補完処理部にデータを送りステップS113からの処理が行われる。もし含まれていない場合は第1の実施形態と同様な処理を行う。上記フローチャートは、図30のようになる。ステップS130で第1の実施形態と同様な処理を行い、サーバ3からの応答をステップS131で属性をチェックする。特殊な属性であれば第2の実施形態のステップS113へ移り、特殊な属性でなければステップS132で第1の実施形態と同様な処理を行う。

【0142】このように、本実施形態では、スイッチ100は、クライアント1からの要求パケット受信時に結果的に行わなくてもよい特殊オブジェクトID管理テーブル127の検索を行わないため、高速な処理が可能になる。その代わりにステップS131でサーバ3からの応答パケットに特殊な属性が含まれているかを判定する必要があるが、特殊オブジェクトID管理テーブル127が複数のエントリがありえることに対してこの判定は応答パケットの特定位置を特殊な属性と比較するだけであり、データ長もオブジェクトIDがNFSであれば32バイトの比較が最低必要なのに対して実装次第で短くすることが可能である。よってステップS111とステップS131を比較するとステップS131は高速な処理が可能である。

【0143】(第9の実施形態)以下では、図15、図31および図32を参照して、本発明の第9の実施形態を詳細に説明する。サーバL(310)およびサーバM(311)はマルチプロトコル対応サーバであり、ネットワークを介したファイルシステムのサービスを行うプロトコルとして、プロトコルAとプロトコルBを解釈することができる。また、プロトコルAは、本発明によるスイッチ装置およびサーバAも対応しているとする。図31に示したように、本発明以外にも、同様の目的を達成するため、クライアントに改造を施す、あるいはサーバに分散ファイルシステムを導入するなどして、プロトコルBによるアクセスに対して他の統合ディレクトリツリーのイメージ320を提供している事例がある。このような環境下において、本発明のスイッチ装置は、プロトコルB向けの統合ディレクトリツリーのイメージ320を取り込み、プロトコルAによるアクセスに対しても

同様の統合ディレクトリツリーのイメージ320を提供することができる。

【0144】具体的には、擬似ファイルシステム105内のエージェントが、統合ディレクトリツリーのイメージ320の情報を入手する。次に、管理者の指示により、統合ディレクトリツリーのイメージ320のルートであるディレクトリAを、スイッチ装置上の擬似ファイルシステムディレクトリツリー200のどこに配置するかを決定する。この例では、擬似ファイルシステムディレクトリツリー200のディレクトリg（すなわち、サーバA（250）のディレクトリg）の下にディレクトリAが見えるように設定する場合を説明する。

【0145】この場合、スイッチ装置はサーバAのディレクトリgの下にダミーディレクトリA'（302）を作成し、さらにサーバLにプロトコルAでアクセスして、そのディレクトリAの下にダミーディレクトリB'（303）を作成する。続いて、特殊オブジェクトID管理テーブル127の内容を、これら2箇所のファイルシステムのつなぎ目がシームレスにクライアントからアクセスできるよう、ディレクトリgとA、ディレクトリAとBのつなぎ目の情報を、たとえば図32に記したように登録する。以上により、上述のように、つなぎ目に関係するアクセス要求については、ファイルシステム補完処理部104は要求を分割し適切なサーバに送付して、その応答より必要な情報を集めて、クライアントに応答を返すので、擬似ファイルシステムディレクトリツリー200のイメージ通りの統合ツリーが提供できる。

【0146】（第10の実施形態）次に、本発明の第10の実施形態のスイッチ装置について説明する。本実施形態のスイッチ装置は、第1、第2および第3の実施形態と同様に、図1に示すネットワークファイルシステムに適用される。本実施形態は、複数のサーバ3間でのオブジェクト群の移動操作を、クライアント1からの移動対象オブジェクト群へのファイルアクセスを停止させずに行うことを目的としている。すなわち第1～第3の実施形態のスイッチ装置では、複数のサーバを統合管理してクライアントにシングルシステムイメージのファイルサービスを提供することができるが、ユーザがサーバ上のオブジェクト（ファイル）に実際にアクセス（書き込みや読み出し）を行っている最中にオブジェクトの再配置を行おうとすると、そのアクセスの期間中はアクセスしようとした際に決定したオブジェクトIDを使い続ける形態となっているので、ユーザのアクセスが中断したり、適切な再配置を行えないことになる。そこでこの第10の実施形態では、ユーザが現にサーバ上のオブジェクトにアクセスしているタイミングであっても管理者側ではオブジェクトの再配置を行うことができ、しかも、オブジェクトの再配置などが行われたことをユーザに隠蔽できるスイッチ装置について説明する。

【0147】この実施形態では、第3の実施形態に関連

して図15を用いて説明したスイッチ装置を使用する。すなわちスイッチ100は、第1、第2、第3の実施形態のスイッチ装置に備えられているオブジェクトID書き換え部101と、ファイルアクセス管理部102と、バケット処理部103と、第2、第3の実施形態に備えられているファイルシステム補完処理部104と、擬似ファイルシステム105とに加え、第3の実施形態に備えられているデータ移動処理部106を備えている。

【0148】図33は、本実施形態におけるファイルアクセス管理部102を示すブロック図である。ファイルアクセス管理部102は、第1、第2、第3の実施形態のスイッチ装置において備えられているトランザクションID管理テーブル125およびサーバ情報管理テーブル126と、第2、第3の実施形態において備えられている特殊オブジェクトID管理テーブル127に加えて、移動処理を行うオブジェクト群の管理を行う移動オブジェクト状態管理テーブル130を備えている。

【0149】次に、本実施形態のスイッチ装置の動作について説明する。本実施形態のスイッチ装置は、第1、第2、第3の実施形態のスイッチ装置において備えられていた機能に加え、擬似ファイルシステム105上に構成されている統合ディレクトリツリーを構成している、各サーバ3のオブジェクトとディレクトリツリー構造を別のサーバ3へコピーをして、移動処理を行っている際に、クライアント1からの移動処理対象オブジェクトへの読み出しや書き込みなどのファイルアクセス要求の転送処理を、オブジェクトの移動処理状態によって変化させることにより、データ移動処理中であることをクライアント1に隠蔽する。

【0150】例えば、第3の実施形態の例で挙げたように、サーバB（251）のディレクトリc（222）からのディレクトリツリーにおけるオブジェクト群のデータと、オブジェクトの所有者、アクセス権などの属性情報を、ストレージ容量の不足といった理由により、ツリー構造ごとサーバA（250）へコピーするのが望ましい状況であったとする。オブジェクトの移動は、前記オブジェクト群のデータをコピーして、図12の擬似ファイルシステム105のマッピング情報を更新し、サーバB（251）からコピーを行った部分を削除することで終了するが、オブジェクトのコピー中に、データの更新などのファイルアクセス操作を要求された際、データの不一致が生じないように、クライアント1からのファイルアクセスをスイッチ100で制御することで、ファイルアクセスを停止することのない、移動処理を実行することが可能となる。

【0151】次に、データの移動処理中の操作に関するスイッチ100での動作について説明する。データ移動処理部106は、すべてのサーバ3における記憶装置5の残り容量データの監視と、スイッチ100で処理されているファイルアクセス要求バケットや応答バケットに

関する履歴データを記録している。そして、データ移動処理部106は、あらかじめ設定しておいた基準よりも、残り容量が少なくなった場合や、特定のサーバ3のみにファイルアクセス要求が集中していた場合に、擬似ファイルシステム105に登録されているデータを参照し、移動元のサーバ3のディレクトリツリーと、移動先のサーバ3の移動場所を決定する。

【0152】データのコピーを実行する前に、データ移動処理部106は、移動元のサーバ3から、移動元のディレクトリツリーのオブジェクト群のディレクトリツリーの構成情報と、すべての元オブジェクトID150を抽出し、前記元オブジェクトID150は、オブジェクトID書き換え部101において、圧縮型情報付帯オブジェクトID151、および追記型情報付帯オブジェクトIDに変換した後、ファイルアクセス管理部102の移動オブジェクト状態管理テーブル130へ登録する。

【0153】移動オブジェクト状態管理テーブル130は、図34のように、移動元の圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155が登録されている移動元オブジェクトID166と、移動先の追記型情報付帯オブジェクトID151と追記型情報付帯オブジェクトID155が登録される移動先オブジェクトID167、移動処理の進行状態の情報を表す進行状態情報177と、対象オブジェクトのディレクトリツリー上の親オブジェクトに相当するオブジェクトの移動オブジェクト状態管理テーブル130上の付与番号を情報としてもつ、親オブジェクト番号178、コピー動作中のクライアント1からの更新履歴情報を登録する更新ログ情報179、対象オブジェクトがディレクトリであるかファイルであるかを示すオブジェクト属性情報180、移動先のサーバ3に対応したサーバ識別情報154を登録する移動先サーバ情報168からなるエントリにより構成されている。

【0154】進行状態情報177へは、コピー処理が行われていないことを表す、未コピー状態、実際にコピー処理を行っている最中であることを表す、コピー中状態、コピーが終了したことを表す、コピー終了状態、移動処理が終了したことを示す、移動終了状態が登録され、各処理の状態により、未コピー状態→コピー中状態→コピー終了状態→移動終了状態という順番で登録が変更される。

【0155】したがって、データ移動処理部106によって、データコピーが実行される前に、移動対象となったすべてのオブジェクトに対し、移動元オブジェクトID166と、ディレクトリ構成情報として、親オブジェクト番号178、およびオブジェクト属性情報180と、移動先サーバ情報168が登録されるとともに、進行状態情報177が、未コピー状態にセットされる。また、移動処理を行うオブジェクト群のディレクトリツリー上、最上位に位置するディレクトリに相当するオブジ

ジェクトのエントリの親オブジェクト番号178は0を登録する。

【0156】ここで、この実施形態の処理の概要を図35を用いて説明する。まず、移動対象のオブジェクトを決定するとともに、移動対象となる全てのオブジェクトに関する情報を移動オブジェクト状態管理テーブル130にマップとして登録する。次に、オブジェクトの移動を開始し、それに伴って、マップ（移動オブジェクト状態管理テーブル130）の更新を行い、移動の終了を確認する。その際、オブジェクトを小さな移動処理単位に分割し、移動処理単位を1つずつ移動元から移動先にコピーするものとし、コピー後、コピーされたオブジェクトの同期を確認するとともに、上述した擬似ファイルシステムにおけるディレクトリの付け替えを行う（FS分岐の作成）。その際、コピー終了領域への更新は、データ移動処理部106によってブロックする。1つずつ移動処理単位をコピーすることを繰り返して、最終的に移動が終了したら、移動元にあるオリジナルのオブジェクトを削除し、これによって処理が完了する。

【0157】次に、データ移動処理部106によって行われるオブジェクト群の移動処理の詳細について、図36のフローチャートを用いて説明する。オブジェクト移動処理部106は、まず、移動オブジェクト状態管理テーブル130へ移動対象となるすべてのオブジェクトに関する情報を登録した後（S140）、親オブジェクト番号178より判別した、ディレクトリ階層上もっとも上位のディレクトリとなるオブジェクトと、親オブジェクト番号178より判別した、前述のオブジェクトの下層に含まれる1階層分のすべてのオブジェクトの移動処理を開始し（S141）、移動元サーバ3から移動先サーバ3へのコピーを実行する（S142）。データ移動処理部106は、実際にコピー処理を行うオブジェクトの移動オブジェクト状態管理テーブル130中の移動先オブジェクトID167へ、移動先のサーバ3から抽出した元オブジェクトID150を、オブジェクトID書き換え部101により、圧縮型情報付帯オブジェクトID151、および追記型情報付帯オブジェクトID155に変換したものを登録後、進行状態情報177がコピー中状態にセットし、移動元から移動先へのコピーを開始する。処理移動先へのコピー処理の終了とともに、進行状態情報177は、コピー終了状態にセットされる。進行状態情報177が、コピー中状態であるオブジェクトへクライアント1からの更新要求が発生した場合、データ解析部124は、更新データ以外の書き込み位置情報などを履歴情報として、更新ログ情報179に追記する。当初のコピー処理が終了した後、更新ログ情報179に情報が登録されているかを確認し（S143）、登録されていれば、更新ログ情報179より更新差分のみのコピー処理を実行して、更新ログ情報179をクリアした後、進行状態情報177をコピー終了状態へセッ

トする(S144)。

【0158】一方、ステップS143において更新ログ情報が登録されていない場合には、1階層分のオブジェクトの進行状態情報177がすべてコピー終了状態にセットされた後、データ移動処理部106は、擬似ファイルシステム105に登録されている統合ディレクトリツリー情報のうち、コピーが終了した部分の移動元のオブジェクト部分のマッピングを移動先のオブジェクトへのマッピングとして変更を加えた後、マッピングを変更したオブジェクトの進行状態情報177を移動終了状態へセットする(S145)。そして前述の最上位ディレクトリに相当するオブジェクトの下層に含まれる1階層分のオブジェクトの中に、オブジェクト属性情報180に登録されている情報がディレクトリの属性であるオブジェクトが存在する、つまりサブディレクトリとなるオブジェクトが存在するかを確認し(S146)、サブディレクトリが存在する場合、データ移動処理部106は、そのサブディレクトリとなるオブジェクトから1階層分のすべてのオブジェクトを、最上位ディレクトリの場合と同様にして、コピー処理を行い(S147)、擬似ディレクトリ105のマッピング情報を変更する。サブディレクトリが存在しない場合、すなわち全ての処理が終了している場合には、ステップS148に移行して、移動元のデータを削除する。

【0159】このようにして、サブディレクトリの有無を確認し、1階層分ずつディレクトリごとにディレクトリツリーをたどって移動処理を実行し、移動オブジェクト状態管理テーブル130に登録されたすべてのオブジェクトの進行状態情報177が、移動終了状態にセットされた後、移動元のサーバ3に存在する移動対象となったオブジェクトをすべて削除して、すべての移動処理を終了する(S148)。

【0160】次に、データ移動処理部106により移動処理が行われている時のクライアント1からのファイルアクセス要求バケット、および応答バケットのスイッチ100での処理方法の詳細について説明する。クライアント1からのファイルアクセス要求は、データや属性の読み出しを実行する読み出し操作、データの書き込みや属性変更などを実行する更新操作、オブジェクトを新しく作成する新規作成操作、オブジェクトを削除する削除操作があり、スイッチ100は各操作によって処理方法を変更している。各操作におけるスイッチ100での処理方法を図37から図41までのフローチャートを用いて説明する。

【0161】クライアント1のファイルアクセス要求バケットを受信すると、バケット処理部103は、ファイルアクセス要求データと、クライアント1のクライアントIPアドレス情報160およびクライアントMACアドレス情報161を要求バケットから抽出し、それらをファイルアクセス管理部102へ送信する(S15

0)。

【0162】データ解析部124は、受信した要求データを解析し、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155に含まれるサーバ識別情報154を抽出し、移動オブジェクト状態管理部130の移動先サーバ情報168に一致するサーバ識別情報154がないか確認する(S151)。

【0163】前述したサーバ識別情報154が移動先サーバ情報168に存在しない場合、以降のスイッチ100における操作は第2の実施形態における要求バケット、および応答バケットの処理と同様の方法とる(S152)。

【0164】一方、ステップS151においてサーバ識別情報154が移動先サーバ情報168に存在する場合は、移動オブジェクト状態管理テーブル130の移動元オブジェクトID166に圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155が登録されているか、エントリを検索して確認する(S153)。

【0165】ここで、移動元オブジェクトID166に、一致する圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトIDがない場合、以降のスイッチ100における操作は第2の実施形態における要求バケット、および応答バケットの処理と同様の方法をとる(S154)。一方、移動元オブジェクトID166に圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトIDが存在する場合、データ解析部124は、要求データからどのような操作要求であるかを判別する(S155)。

【0166】ステップS155において、要求データから得られた操作要求が読み出し操作であった場合(S160)には、移動オブジェクト状態管理テーブル130の該当する移動元オブジェクトID166のエントリの進行状態情報177を参照し、未コピー状態、コピー中状態、コピー済み状態のいずれかの状態であった場合(S161)、移動元のサーバ3へ要求バケットを転送することになり(S162)、以降のスイッチ100における操作は、第2の実施形態における要求バケット、および応答バケットの処理と同様の方法をとる(S163)。

【0167】ステップS161において、未コピー、コピー中、コピー済みのいずれでもない場合は、移動済み状態であった場合(S164)であり、移動先のサーバ3へ要求バケットを転送することになり(S165)、移動オブジェクト状態管理テーブル130の該当した移動元オブジェクトID166のエントリの移動先オブジェクトID167に登録されている圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155に変換し(S166)、以降のスイッチ100における操作は、第2の実施形態における要求バ

ケット、および応答ケットの処理と同様の方法をとる（S163）。

【0168】ステップS155において要求データから得られた操作要求が更新操作であった場合（S170）、移動オブジェクト状態管理テーブル130の該当する移動元オブジェクトID166のエントリの進行状態情報177を参照し、未コピー状態であった場合（S171）、移動元のサーバ3へ要求ケットを転送することになり（S172）、以降のスイッチ100における操作は、第2の実施形態における要求ケット、および応答ケットの処理と同様の方法をとる（S173）。

【0169】ステップS171において未コピー状態でない場合には、進行状態情報177から操作対象がコピー中かどうかを判断し、コピー中状態であった場合（S174）、移動元サーバ3へ要求ケットを転送することになり（S175）、その該当エントリの更新ログ情報179に、書き込み位置情報などの更新データ以外のメタ情報を追記し（S176）、以降のスイッチ100における操作は、第2の実施形態における要求ケット、および応答ケットの処理と同様に方法をとる。ただし、移動元サーバ3からの応答データが、クライアント1に削除権限がないなどの理由でエラーとなっている場合は、応答データをクライアント1へ転送し、移動オブジェクト状態管理テーブル130の更新ログ情報179へ登録したデータを削除する（S177）。

【0170】ステップS174においてコピー中でない場合には、進行状態情報177を参照して操作対象がコピー済みであるかどうかを判断し、コピー終了状態であった場合（S178）、移動元サーバ3および移動先サーバ3の両方へ転送することになり（S179）、トランザクションID管理テーブル125に必要な情報を登録した後、要求ケットに含まれる圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155をオブジェクトID書き換え部101で変換した元オブジェクトID150と、圧縮型情報付帯オブジェクトID151もしくは追記型情報付帯オブジェクトID155から抽出したサーバ識別情報154を元に、サーバ情報管理テーブル126から抽出した移動元サーバ3のIPアドレスやMACアドレス情報などの転送先のアドレス情報をファイルシステム補完処理部104に送る。

【0171】同様に、該当移動元オブジェクトID166のエントリの移動先オブジェクトID167に登録されている、圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155をオブジェクトID書き換え部101で変換した元オブジェクトID150と、移動先サーバ情報168に登録されているサーバ識別情報154を元に、サーバ情報管理テーブル126から抽出した移動先サーバ3のIPアド

レスやMACアドレス情報などの転送先のアドレス情報をファイルシステム補完処理部104に送る。

【0172】ファイルシステム補完処理部104では、ファイルアクセス管理部102より受け取った、移動元サーバ3向けのデータと、移動先サーバ3向けのデータを元に、要求ケットを移動元サーバ3、および移動先サーバ3へパケット処理部103を介してスイッチ100から送信する（S180）。

【0173】移動元サーバ3、および移動先サーバ3からのそれぞれの応答ケットをスイッチ100で受け取り、パケット処理部103でそれぞれの応答データを抽出して、ファイルシステム補完処理部104に送り、ファイルシステム補完処理部104において、2つの応答データを受け取ったことを確認した後、応答データをファイルアクセス管理部102へ送る。以降のスイッチ100での操作は、第1の実施形態の応答ケットの処理方法と同様な方法を取り、クライアント1へ応答ケットが返される（S181）。

【0174】ステップS178においてコピー済みでない場合は、操作対象が移動済み状態であるので（S182）、移動先のサーバ3へ要求ケットを転送することになり（S183）、移動オブジェクト状態管理テーブル130の該当した移動元オブジェクトID166のエントリの移動先オブジェクトID167に登録されている圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155に変換し（S184）、以降のスイッチ100における操作は、第2の実施形態における要求ケット、および応答ケットの処理と同様の方法をとる（S185）。

【0175】ステップS155において要求データから得られた操作要求が新規作成操作であった場合（S190）、移動オブジェクト状態管理テーブル130の新規作成を行うディレクトリに相当するオブジェクトに該当する移動元オブジェクトID166のエントリの進行状態情報177を参照し、未コピー状態、コピー中状態、コピー済み状態のいずれかであるかを判定し（S191）、未コピー状態、コピー中状態、コピー済み状態のいずれかである場合には、移動元のサーバ3へ要求ケットを転送することになり（S192）、以降のスイッチ100におけるパケット転送操作は、第2の実施形態における要求ケット、および応答ケットの処理と同様の方法をとる（S193）とともに、データ移動処理部106は移動オブジェクト状態管理テーブル130へ作成されたオブジェクトのエントリを加える（S194）。

【0176】ステップ190において未コピー状態、コピー中状態、コピー済み状態のいずれでもない場合、すなわち移動済み状態であった場合（S195）、移動先のサーバ3へ要求ケットを転送することになり（S196）、移動オブジェクト状態管理テーブル130の該

当した移動元オブジェクトID166のエントリの移動先オブジェクトID167に登録されている圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155に変換し(S197)、以降のスイッチ100における操作は、第2の実施形態における要求パケット、および応答パケットの処理と同様の方法をとる(S198)。

【0177】ステップS155において 要求データから得られた操作要求が削除操作であった場合(S200)、移動オブジェクト状態管理テーブル130の該当する移動元オブジェクトID166のエントリの進行状態情報177を参照し、未コピー状態であるかどうかを判定し(S201)、未コピー状態であった場合には、移動元のサーバ3へ要求パケットを転送することになり(S202)、以降のスイッチ100におけるパケット転送操作は、第2実施の形態における要求パケット、および応答パケットの処理と同様の方法をとる(S203)とともに、データ移動処理部106は、移動オブジェクト状態管理テーブル130に登録されている削除操作されるオブジェクトのエントリを削除する。ただし、移動元サーバ3から返答される応答データが、クライアント1に削除権限がないなどの理由でエラーとなっている場合は、応答データをクライアント1へ転送し、該当するオブジェクトの移動オブジェクト状態管理テーブル130からの削除は行わない(S204)。

【0178】ステップS201において未コピー状態でない場合には、進行状態情報177によりコピー中状態であるかどうかを判断し(S205)、コピー中状態であれば、移動元のサーバ3、および移動先のサーバ3へ要求を送信することになり(S206)、データ移動処理部106へ該当するオブジェクトのコピー処理を停止させる(S207)。書き込み操作の場合と同様にして、移動元サーバ3、および移動先サーバ3へ要求パケットの送信と、それぞれから返答される応答パケット処理を実行(S208)した後、データ移動処理部106は、該当するオブジェクトの移動オブジェクト状態管理テーブル130のエントリを削除する。ただし、移動元サーバ3および移動先サーバ3から返答される応答データが、クライアント1に削除権限がないなどの理由によりエラーとなった場合は、応答データをクライアント1へ転送するとともに、該当するオブジェクトの移動オブジェクト状態管理テーブル130からの削除は行わずに、データ移動処理部106での該当するオブジェクトのコピーを再開させる(S209)。

【0179】ステップS205においてコピー中状態でない場合には、進行状態情報177によりコピー終了状態であるかどうかを判断し(ステップS210)、コピー終了状態である場合には、移動元のサーバ3および移動先のサーバ3に転送することになり(S211)、コピー中状態の場合と同様にして、ステップS180か～

S171と同様の、移動元サーバ3、および移動先サーバ3へ要求パケットの送信と、それぞれから返答される応答パケット処理を実行した後(S212)、ステップS204に移行し、データ移動処理部106は、該当するオブジェクトの移動オブジェクト状態管理テーブル130のエントリを削除する。ただし、移動元サーバ3および移動先サーバ3から返答される応答データが、クライアント1に削除権限がないなどの理由でエラーとなっている場合は、応答データをクライアント1へ転送し、該当するオブジェクトの移動オブジェクト状態管理テーブル130からの削除は行わない。

【0180】ただし、要求データから得られた操作対象のオブジェクトがディレクトリであり、かつ進行処理状態情報177がコピー中、もしくはコピー終了状態であった場合、移動元サーバ3と移動先サーバ3への要求送信の方法は、同時に要求を送信するのではなく、まず移動元サーバ3へ削除要求を送信し、削除操作要求が成功した場合のみ、移動先サーバに要求を送信することとし、移動元サーバ3においてエラーとなった場合は、移動先サーバ3へは削除要求を送らない。

【0181】ステップS210においてコピー済み状態でない場合は、操作対象が移動済み状態である場合であるから(S213)、移動先のサーバ3へ要求パケットを転送することになり(S214)、移動オブジェクト状態管理テーブル130の該当した移動元オブジェクトID166のエントリの移動先オブジェクトID167に登録されている圧縮型情報付帯オブジェクトID151、もしくは追記型情報付帯オブジェクトID155に変換し(S215)、以降のスイッチ100における操作は、第2の実施形態における要求パケット、および応答パケットの処理と同様の方法をとる(S216)。

【0182】データ移動処理部106によるオブジェクト群の移動処理がすべて完了し、移動元サーバ3の該当するオブジェクトが削除された後、移動オブジェクト状態管理テーブル130を用いた、要求パケット、および応答パケット処理は、移動オブジェクト状態管理テーブル130のエントリを一定時間後にデータ移動処理部106によって削除されるまで有効となる。

【0183】(第11の実施形態)本実施形態は、第10の実施形態をさらに変更した実施形態である。第10の実施形態では、移動対象となったすべてのオブジェクトの移動処理の終了後、移動元サーバ3の移動対象オブジェクトを削除していたが、本実施の形態では、前記移動元サーバ3の移動対象オブジェクトの削除を行なわない。また、移動処理中に、クライアント1からの更新操作要求、新規作成要求、削除要求のいずれかの操作要求が来た場合、操作対象オブジェクトの移動オブジェクト状態管理テーブル130の進行状態情報177が移動済み状態であっても、コピー済み状態と同じ処理をスイッチ100で行うことで、移動元の移動対象のオブジェクトと

そのオブジェクトの複製である移動先の移動対象のオブジェクト間で同期を取り続けることが可能となる。

【0184】（第12の実施形態）本実施形態は、第10の実施形態および第11の実施形態をさらに変更した実施形態である。第10の実施形態、もしくは第11の実施形態では、スイッチ100での移動処理中にクライアント1からの更新操作要求が来た場合、操作対象となっているオブジェクトに移動オブジェクト状態管理テーブル130の更新ログ情報179に更新のメタ情報だけでなく、更新データも登録し、後々の更新部分のコピー処理時に、移動元サーバ3から読み出しを行わずに、直接移動先サーバ3の更新操作対象オブジェクトへ更新処理を行う。

【0185】（まとめ）なお、前述のように、各実施形態に示すスイッチ装置は、図1に示すネットワーク構成のネットワークファイルシステムに適用することができるが、各実施形態のスイッチ装置は、図20に示すように、ネットワーク2ではなく、スイッチ100にサーバ3を直接接続する構成であっても、上述の機能を提供することが可能である。このようなネットワーク構成とすれば、クライアント1からサーバ3へのスイッチ100を経由しないアクセスを完全に遮断することができるため、クライアント1から直接サーバ3へ不正にアクセスすることによって発生するシステムダウンを防ぐことが可能となる。

【0186】また、各実施形態に示すスイッチ装置では、オブジェクトID書き換え部101の代わりに、サーバ3が生成した元オブジェクトID150とサーバ識別情報154との組み合わせから成るエントリを有するテーブル（第1のテーブル）を備えていてもよい。この場合、データ解析部124は、このテーブルを参照して、ファイルアクセス要求に含まれるオブジェクトIDに対応するサーバ識別情報を求め、サーバ情報管理テーブル126を参照して、そのサーバ識別情報に対応するサーバのアドレス情報を求め、そのアドレス情報に基づいてそのファイルアクセス要求の転送先を決定する。このようにすれば、元オブジェクトID150をクライアント1へそのまま渡しても同様の機能を実現することができる。しかし、この場合には、同じデータ列を有する元オブジェクトID150が生成される可能性もあるため、そのテーブルに既に登録されている元オブジェクトID150と同一のデータ列を有するオブジェクトIDをスイッチ100で受信した場合は、スイッチ100において、そのオブジェクトIDを別のデータ列に書き換え、前述のテーブルに書き換えられたデータ列を登録するエントリを設けるなどの処理が必要となる。

【0187】また、上述した各実施形態のスイッチ装置は、ネットワークファイルシステムに適用されているが、本発明はこれに限定されるものではなく、複数のサーバが所有するリソース（オブジェクト）をクライアン

トが利用するシステムであれば、ネットワークファイルシステム以外にも適用することができる。

【0188】なお、スイッチ100には、スイッチ100の動作を実行するためのプログラムが記録されている。このプログラムは、コンピュータサーバであるスイッチ100のハードウェアの動作を制御するものであり、上述したスイッチ100の処理は、全てそのプログラムによって指定されているものである。

【0189】

10 【発明の効果】以上述べたように、本発明のスイッチ装置は以下に示す効果を有する。

【0190】（1） クライアントおよびサーバが標準のネットワークを介したファイルシステムサービスプロトコル（例えばNFS）さえサポートしていれば、それらの既存システムに手を加えることなく、複数のサーバを統合管理してクライアントにSSIのファイルサービスを実現できることにある。本発明のスイッチ装置の実装にあたっては、本発明のスイッチ装置を論理的にクライアントとサーバの間に配置するのみで、IPEイリアスの設定や仮想分散ファイルシステムの導入などのクライアントおよびサーバの処理の変更が不要であるからである。

20 【0191】（2） 管理者が負荷や容量分散の目的という本来クライアントに対して通知する必要のない操作、例えばクライアントに一度割り当てたサーバから別のサーバへオブジェクトを移動するような場合に、クライアントに対して透過的に処理を実行することができ、クライアントに提示したSSIの一貫性を維持することができる。本発明のスイッチ装置は、リモートファイルシステムにおいてディレクトリやファイルなどのオブジェクトを特定するためのオブジェクトIDとそのオブジェクトを処理するサーバを対応付けし、オブジェクトをクライアント透過に移動させた場合の新旧オブジェクトIDの変換などを行い、その結果に基づきクライアントからの要求を適切なサーバへ、またサーバからのオブジェクトIDを含む応答を適切なクライアントへ振り分けることができるからである。

30 【0192】（3） 大量のオブジェクトIDを、検索操作を介さずに処理して、パケットを適切なサーバへ転送することが可能であり、高速化、およびハードウェア実装に有利となる。クライアントに渡すオブジェクトIDを、サーバが発行した元オブジェクトIDから情報付帯オブジェクトIDへ書き換えるためである。

40 【0193】（4） 名前解決処理による負担をサーバとスイッチ装置とで分散して、名前解決処理の時間を短縮することができる。本発明によるスイッチ装置は、擬似ファイルシステムを備え、その擬似ファイルシステムがマウントポイントのみを管理し、そのマウントポイント配下の名前解決は矛盾なく各サーバへ委譲できるからである。

【0194】(5) 統合ディレクトリツリーにおける各共有オブジェクト単位のディレクトリツリーのつなぎ目を含むオブジェクトの移動要求があった場合に、つなぎ目以下のオブジェクトの実際の移動を行わず、そのつなぎ目以下のオブジェクトを所望の位置へと移動したと同等にクライアントに見せるようにテーブルを書き換えるので、非常に高速に移動を完了することができる。

【0195】(6) O S I 参照モデルの第4層以下のタグ情報を使用できるようにした場合には、これらのタグ情報を用いてハードウェアによるオブジェクト I D 書き換えが可能となるので、さらに処理を高速に行うことができる。

【0196】(7) さらに、管理者によるサーバ間でのオブジェクト再配置作業時においても、ユーザへは再配置作業を隠蔽させることで、システム停止や、ファイルアクセスサービス停止を発生させずに管理作業を実行することを可能にする。

【図面の簡単な説明】

【図1】本発明の第1の実施形態のスイッチ装置が設けられたクライアント・サーバ形式のリモートファイルシステムであるネットワークファイルシステムのブロック図である。

【図2】本発明の第1の実施形態のスイッチの構成を示すブロック図である。

【図3】ファイルアクセス管理部の構成を示すブロック図である。

【図4】オブジェクト I D 書き換え部の構成を示すブロック図である。

【図5】オブジェクト書き換え部において書き換えられるオブジェクト I D のフォーマットの一例を示す図である。

【図6】アルゴリズムテーブルの構成を示す図である。

【図7】サーバ情報管理テーブル構成を示す図である。

【図8】トランザクション I D 管理テーブルの構成を示す図である。

【図9】応答データを受信した際のオブジェクト書き換え部の動作を示すフローチャートである。

【図10】本発明の第2の実施形態のスイッチ装置の構成を示すブロック図である。

【図11】本発明の第2の実施形態のスイッチ装置におけるファイルアクセス管理部の構成を示すブロック図である。

【図12】擬似ファイルシステム 105 のディレクトリツリー 200 の一例を示す図である。

【図13】特殊オブジェクト I D 管理テーブルの構成を示す図である。

【図14】本発明の第2の実施形態のスイッチ装置の動作を示すフローチャートである。

【図15】本発明の第3の実施形態のスイッチ装置の構成を示すブロック図である。

【図16】本発明の第3の実施形態のスイッチ装置におけるファイルアクセス管理部の構成を示すブロック図である。

【図17】データの移動を示す図である。

【図18】移動オブジェクト I D 管理テーブルの構成を示す図である。

【図19】バケットを受信した場合のスイッチの動作を示すフローチャートである。

【図20】本発明のスイッチ装置が適用されるクライアント・サーバ形式のリモートファイルシステムであるネットワークファイルシステムのブロック図である。

【図21】オブジェクト I D のフォーマットの一例を示す図である。

【図22】オブジェクト I D のフォーマットの一例を示す図である。

【図23】第6の実施形態における、本発明による擬似ファイルシステムの具体的な構成法を示す図である。

【図24】本発明によるスイッチ装置が、通常のファイルシステム処理によりディレクトリ移動を実施した場合の例を示す図である。

【図25】本発明による第6の実施形態の動作を示す図である。

【図26】本発明の第7の実施形態のスイッチ装置におけるファイルアクセス管理部の構成を示すブロック図である。

【図27】処理分類テーブルの構成を示す図である。

【図28】本発明の第7の実施形態のハードウェア構成のブロック図である。

【図29】本発明の第8の実施形態の想定するディレクトリツリー構成の例を示す図である。

【図30】本発明の第8の実施形態の動作を説明するフローチャートである。

【図31】他の方法、プロトコルによる統合ディレクトリツリー環境が並存する例を示す図である。

【図32】本発明の第9の実施形態による、統合ディレクトリツリー環境の取り込みを説明するための図である。

【図33】本発明の第10の実施形態におけるファイルアクセス管理部内の構成を示すブロック図である。

【図34】本発明の第10の実施形態における移動オブジェクト状態管理テーブルの構成を示す図である。

【図35】本発明の第10の実施形態での処理の概要を説明する図である。

【図36】本発明の第10の実施形態におけるデータ移動処理部におけるオブジェクトの移動処理手順をフローチャートで示した図である。

【図37】本発明の第10の実施形態におけるデータ移動処理部の移動処理中のクライアントからの要求バケットのスイッチ装置における処理手順をフローチャートで示した図である。

【図38】読み出し操作要求時の処理手順をフローチャートで示した図である。

【図39】更新操作要求時の処理手順をフローチャートで示した図である。

【図40】新規作成操作要求時の処理手順をフローチャートで示した図である。

【図41】削除操作要求時の処理手順をフローチャートで示した図である。

【図42】ウェブスイッチの機能を示すブロック図である。

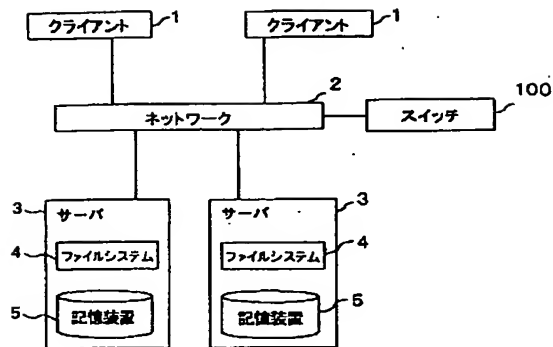
【図43】特開2001-51890号公報に開示された方法を実行するシステムの構成を示すブロック図である。

【符号の説明】

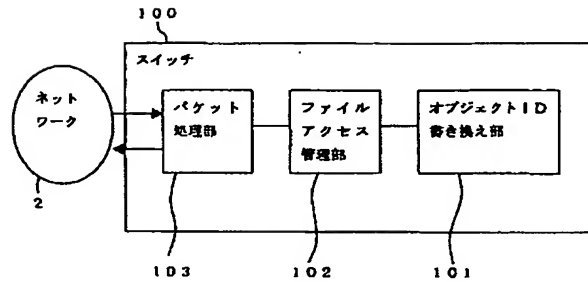
- 1 クライアント
- 2 ネットワーク
- 3 サーバ
- 4 ファイルシステム
- 5 記憶装置
- 6 ローカルファイルシステム
- 7 サーバ間専用結線
- 100 スイッチ
- 101 オブジェクトID書き換え部
- 102 ファイルアクセス管理部
- 103 パケット処理部
- 104 ファイルシステム補完処理部
- 105 擬似ファイルシステム
- 106 データ移動処理部
- 107 データ解析高速処理部
- 108 データ解析非高速処理部
- 120 要求応答分配部
- 121 要求処理部
- 122 応答処理部
- 123 アルゴリズムテーブル
- 124 データ解析部
- 125 トランザクションID管理テーブル
- 126 サーバ情報管理テーブル
- 127 特殊オブジェクトID管理テーブル
- 128 移動オブジェクトID管理テーブル
- 129 処理分類テーブル
- 130 移動オブジェクト状態管理テーブル
- 150 元オブジェクトID
- 151 圧縮型情報付帯オブジェクトID
- 152 圧縮データ
- 153 アルゴリズム識別情報
- 154 サーバ識別情報
- 155 追記型情報付帯オブジェクトID

- 156 削除データ
- 157 削除データ位置
- 158 サーバIPアドレス情報
- 159 サーバMACアドレス情報
- 160 トランザクションID
- 161 クライアントIPアドレス情報
- 162 クライアントMACアドレス情報
- 163 ファイルアクセス命令情報
- 164 オブジェクトID
- 10 165 命令情報
- 166 移動元オブジェクトID
- 167 移動先オブジェクトID
- 168 移動先サーバ情報
- 169 データ移動フラグ
- 170 フラグ付き圧縮型情報付帯オブジェクトID
- 171 フラグ付き追記型情報付帯オブジェクトID
- 172 データ移動世代番号
- 173 世代番号付き圧縮型情報付帯オブジェクトID
- 20 174 世代番号付き追記型情報付帯オブジェクトID
- 175 ポート番号情報
- 176 処理分類情報
- 177 進行状態情報
- 178 親オブジェクト番号
- 179 更新ログ情報
- 180 オブジェクト属性情報
- 200 擬似ファイルシステムディレクトリツリー
- 220 ルートディレクトリ
- 30 221~224、300 ディレクトリ
- 225、301~303 ダミーディレクトリ
- 226~27 オブジェクトグループ
- 250 サーバA
- 251 サーバB
- 310 サーバL
- 311 サーバM
- 320 他の統合ディレクトリツリーのイメージ
- 900 ウェブスイッチ
- 902 HTTP処理部
- 40 903 パケット処理部
- S100~108、S110~S117、S120~S132 ステップ
- S140~148、S150~S155、S160~S166 ステップ
- S170~185、S190~S198、S200~S216 ステップ

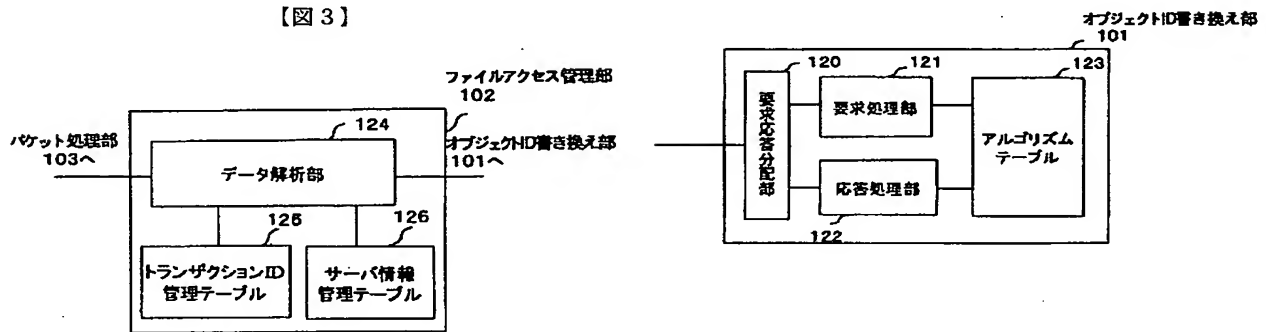
【図1】



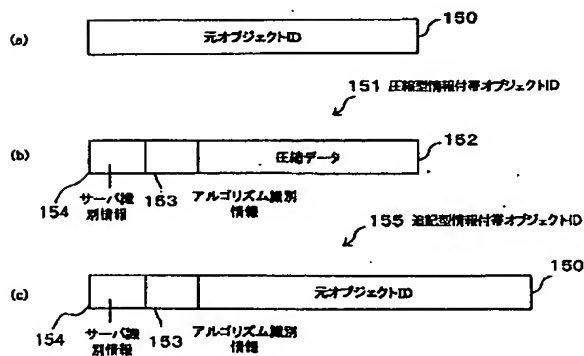
【図2】



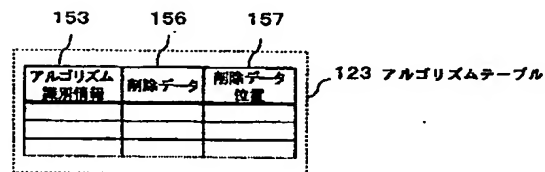
【図4】



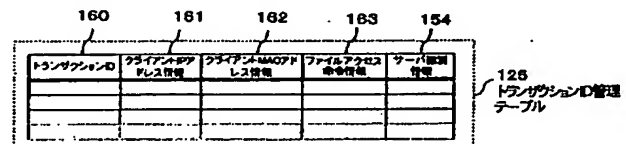
【図5】



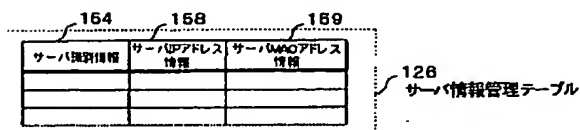
【図6】



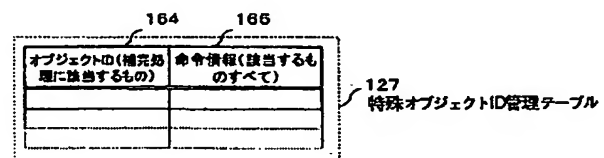
【図8】



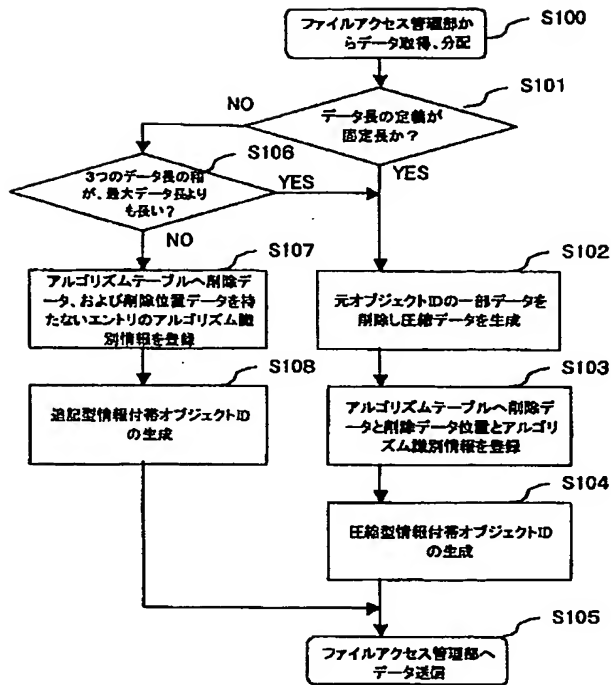
【図7】



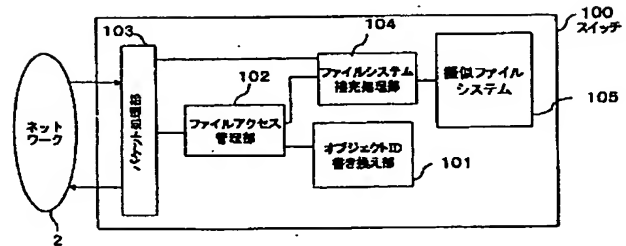
【図13】



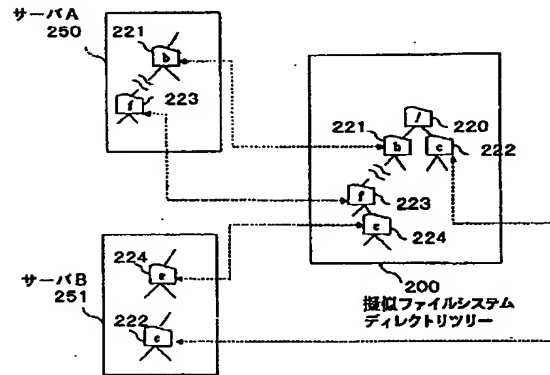
【図9】



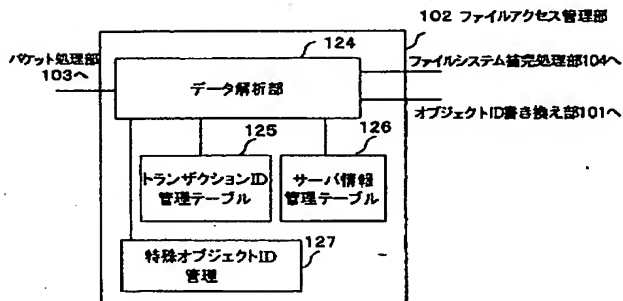
【図10】



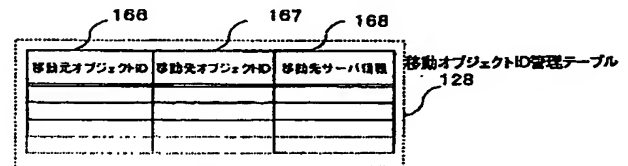
【図12】



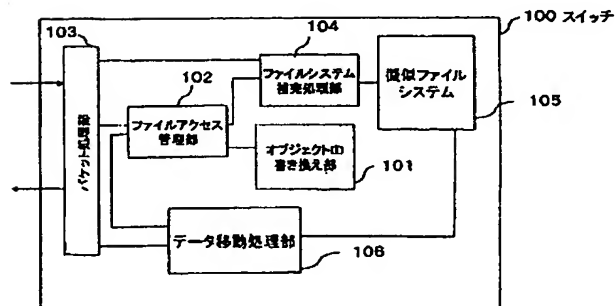
【図11】



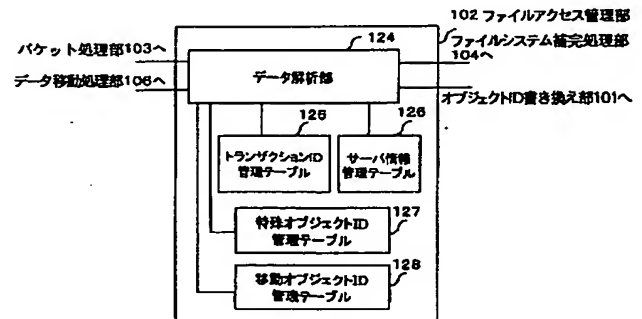
【図18】



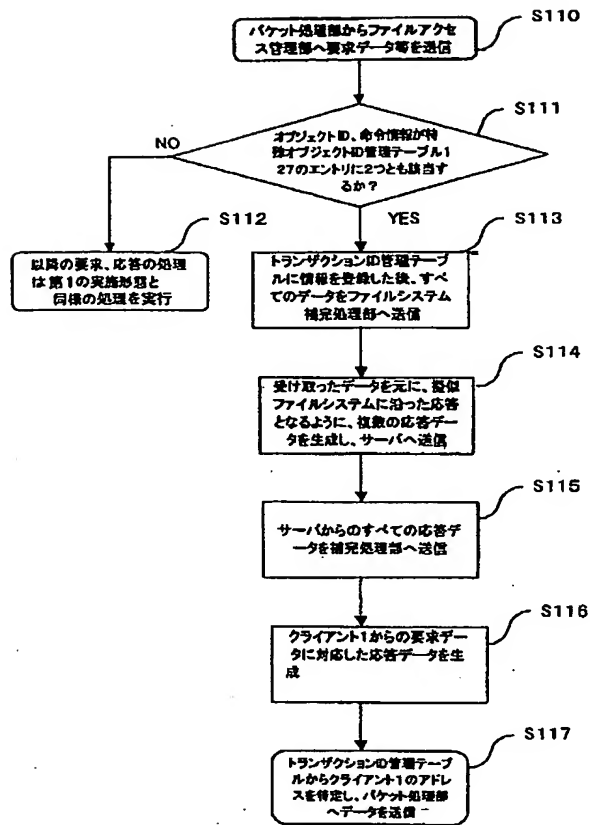
【図15】



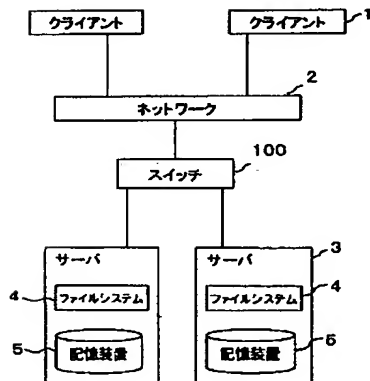
【図16】



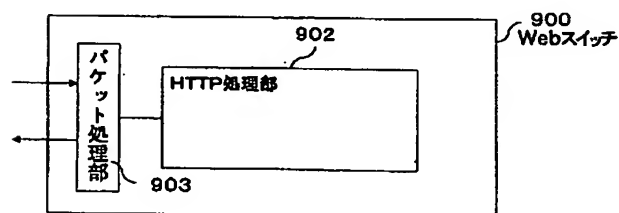
【図14】



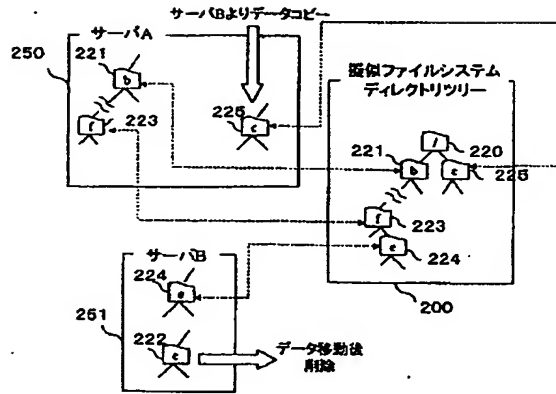
【図20】



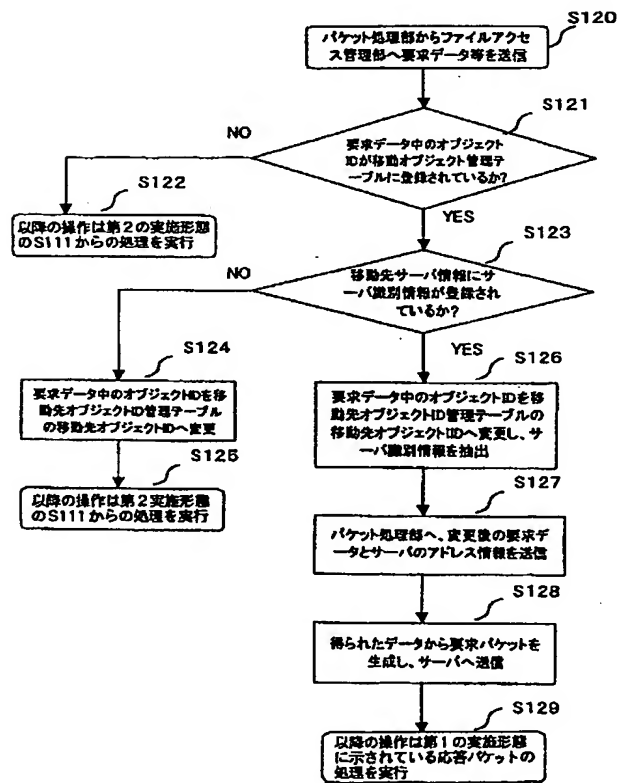
【図42】



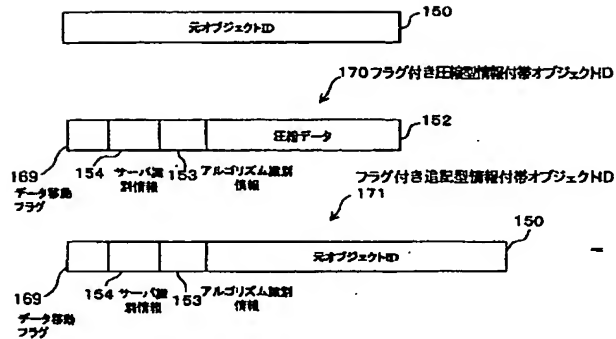
【図17】



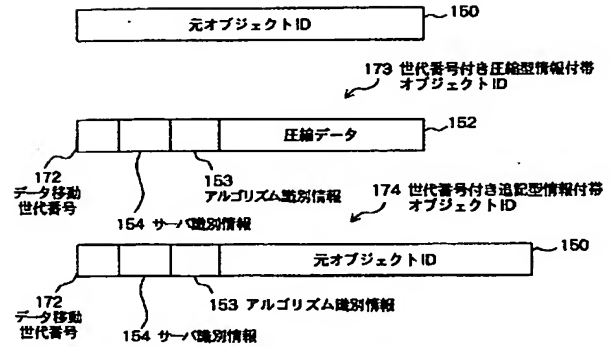
【図19】



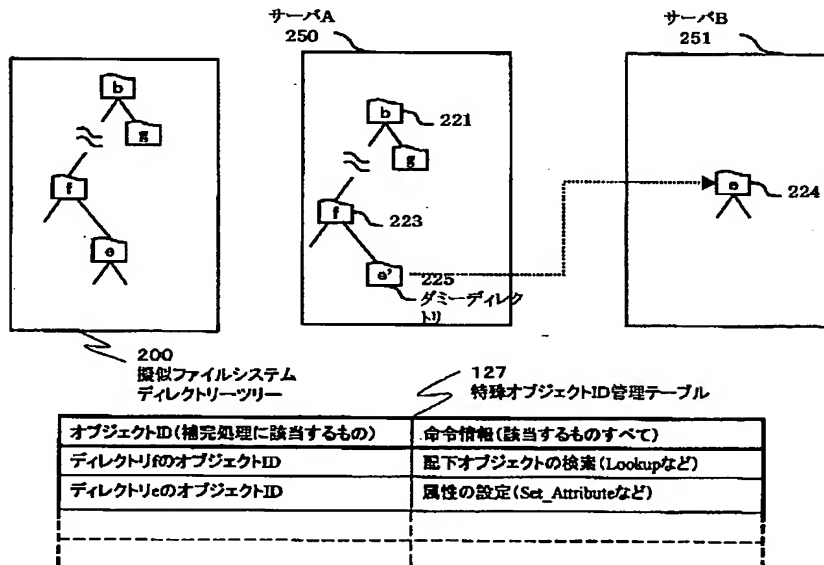
【図21】



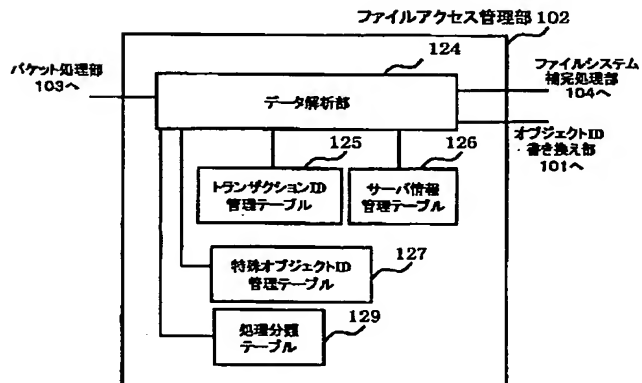
【図22】



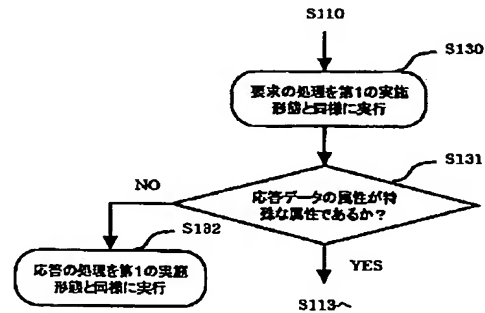
【図23】



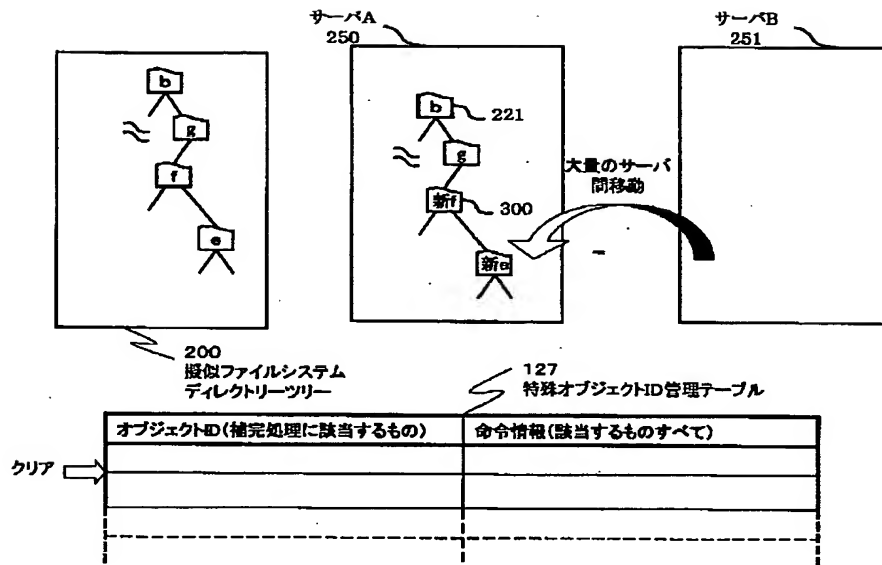
【図26】



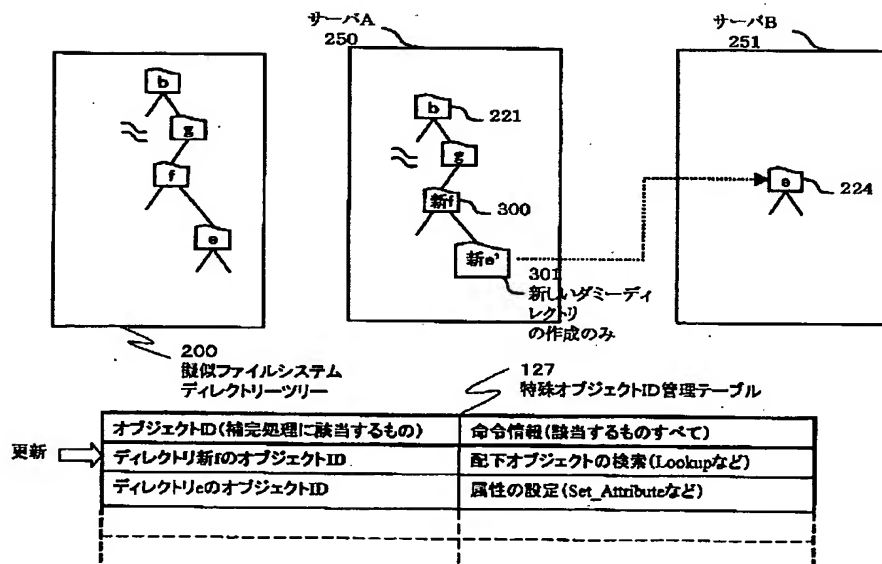
【図30】



【図24】



【図25】

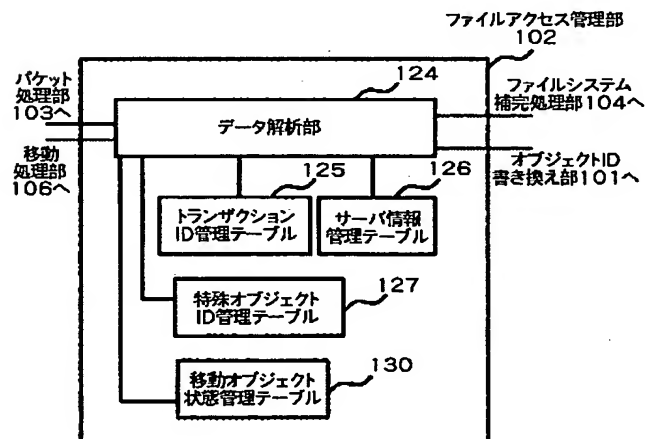


154 サーバ 識別情報	175 ポート番号 情報	161 クライアントIP アドレス情報	162 クライアントMAC アドレス情報	176 処理分類 情報

129 処理分類
テーブル

[illegible]

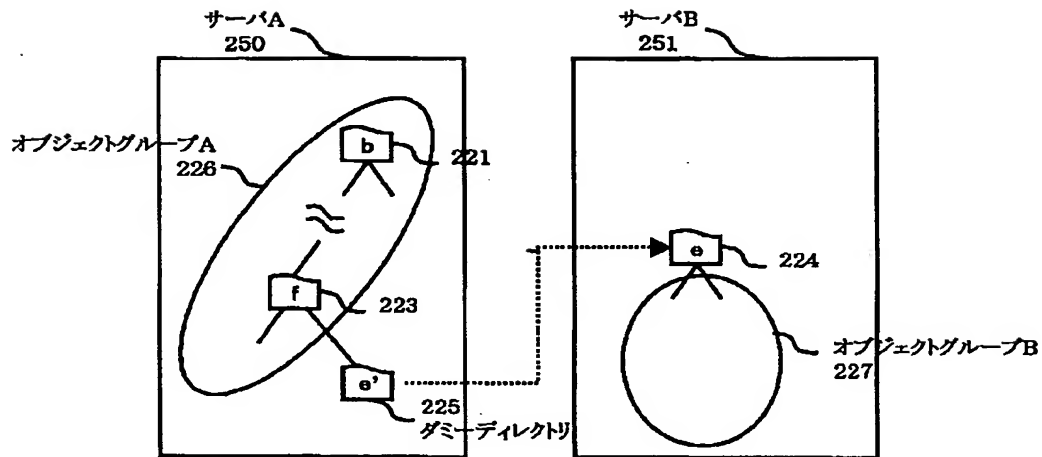
【図 3 4】



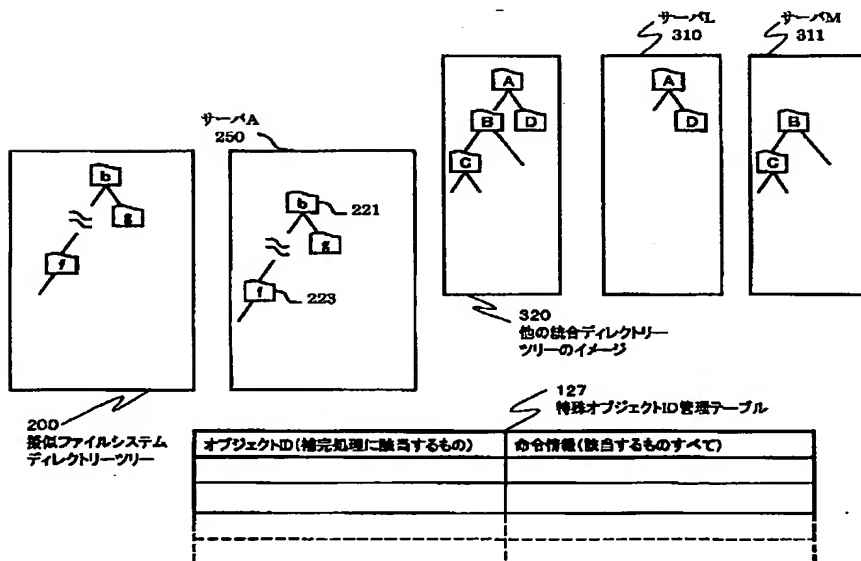
	166	167	177	178	179	180	168
番号	移動元 オブジェクト ID	移動先 オブジェクト ID	進行状態 情報	誤オブジェ クト番号	更新ログ 情報	オブジェクト 属性情報	移動先 サーバ情報
1							
2							
3							
4							

移動オブジェクト状態管理テーブル 130

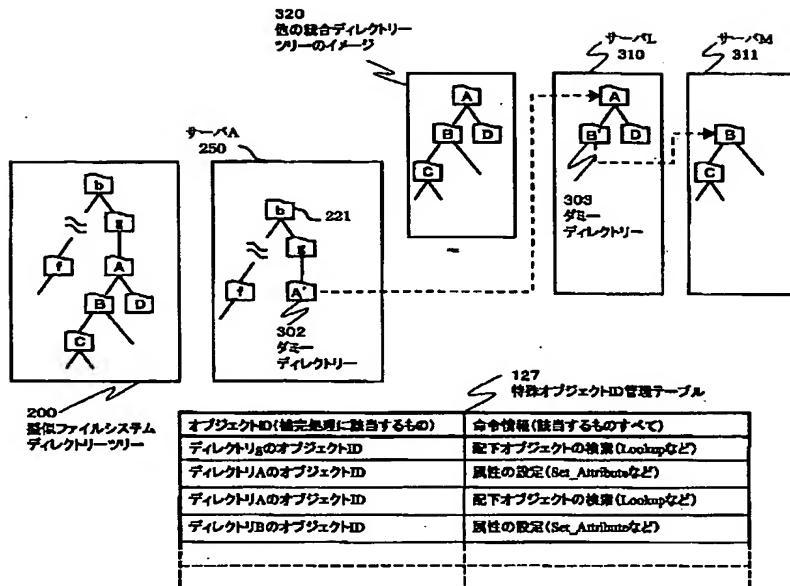
【図29】



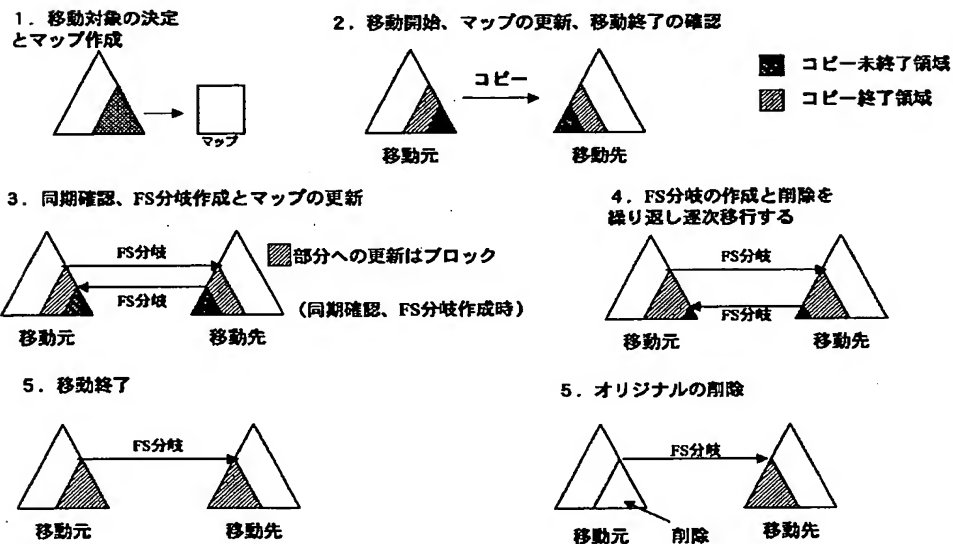
【図31】



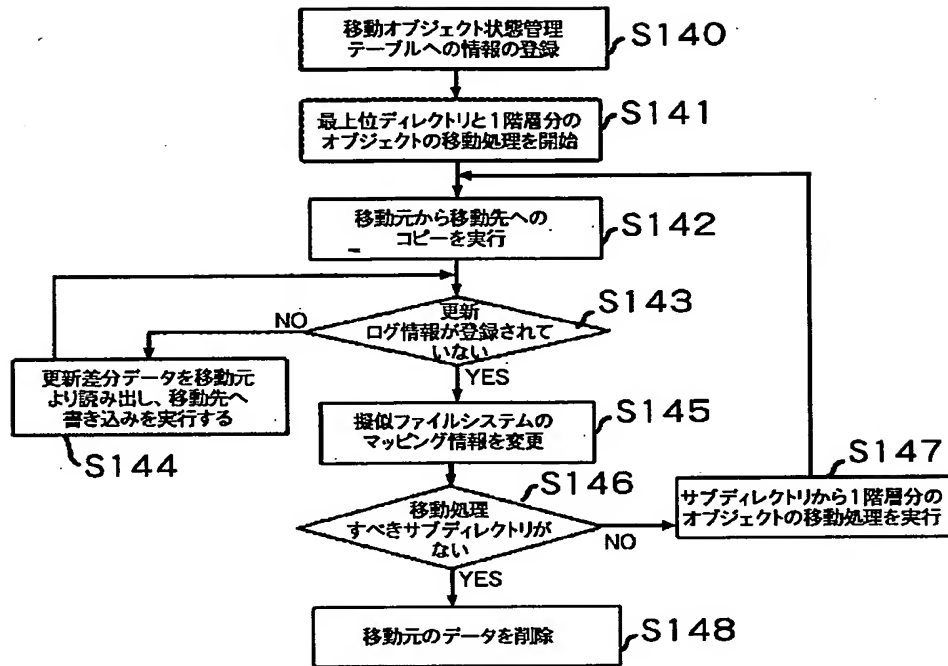
【図32】



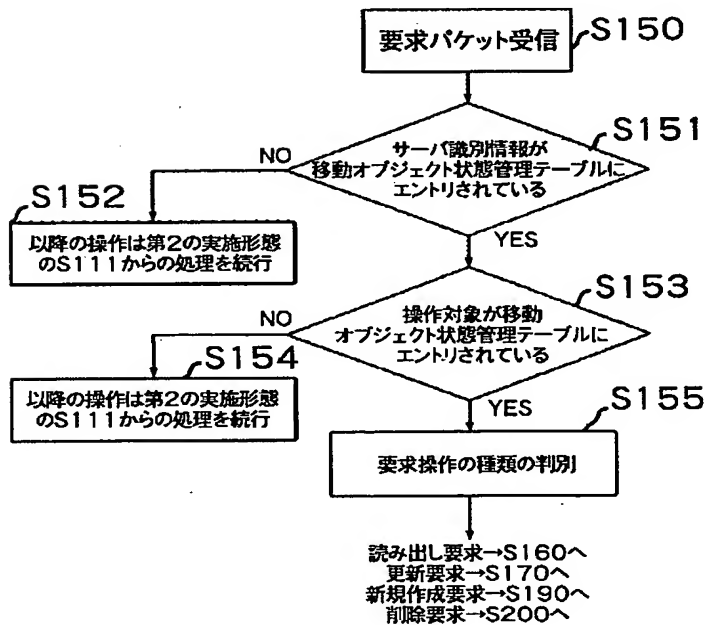
【図35】



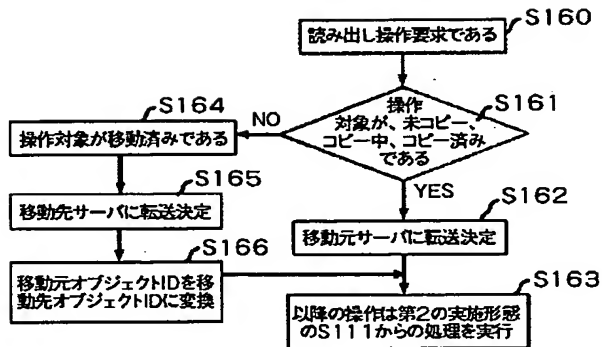
【図36】



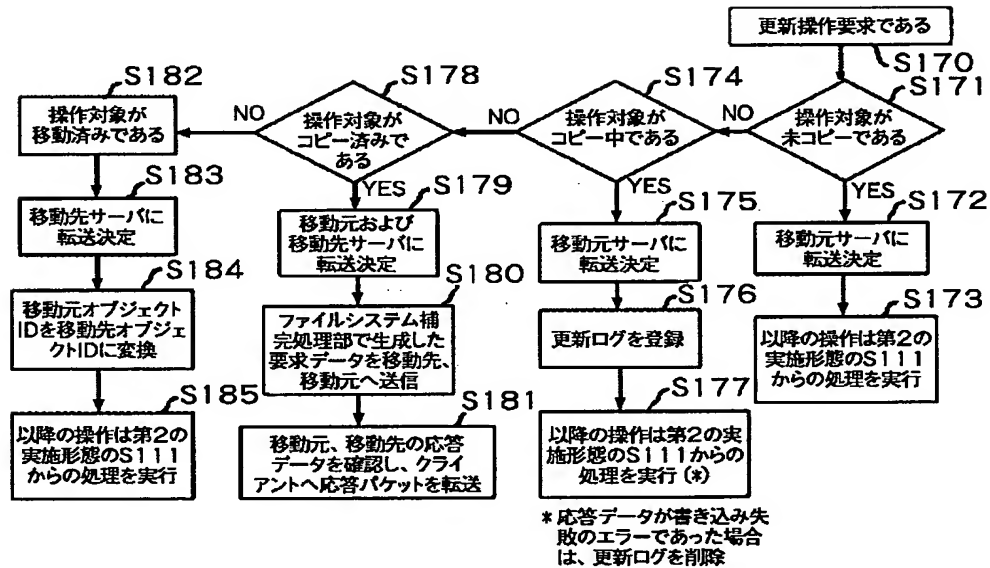
【図37】



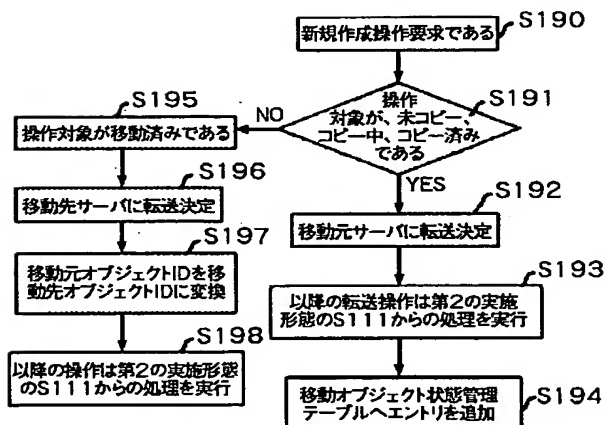
【図38】



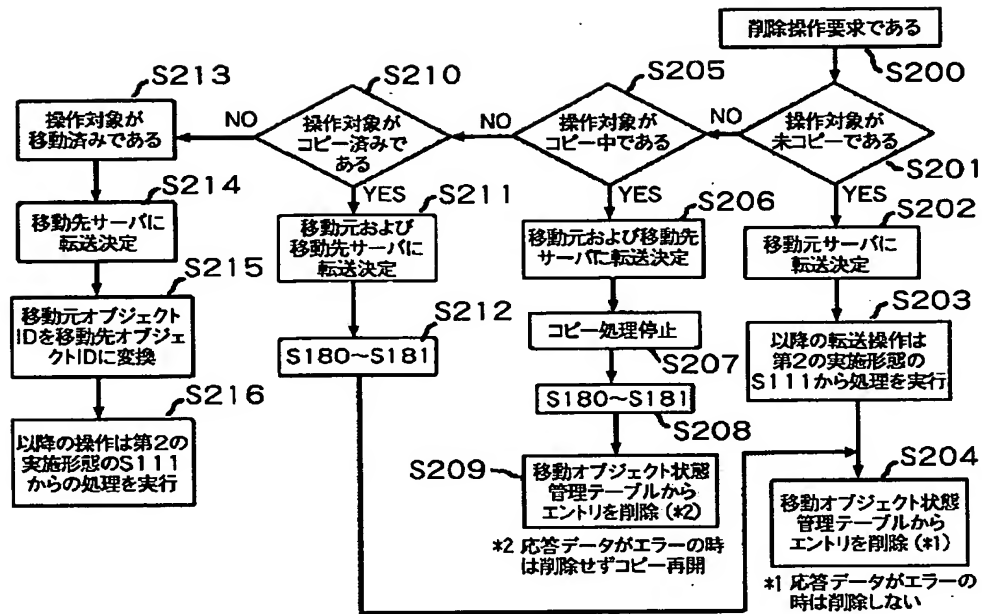
【図39】



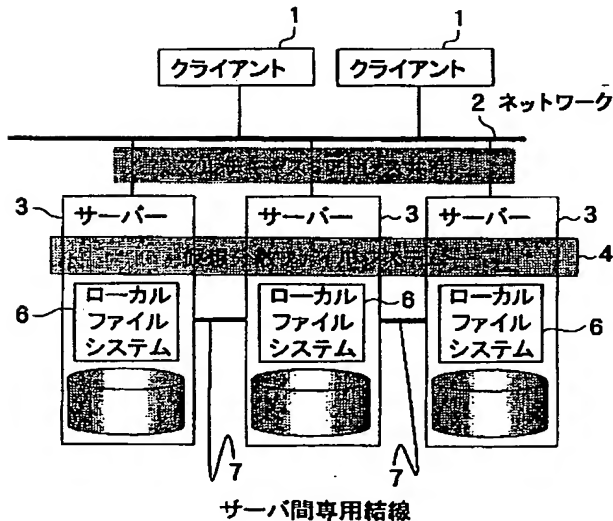
【図40】



【図41】



【図43】



フロントページの続き

(72)発明者 鳥居 隆史
東京都港区芝五丁目7番1号 日本電気株
式会社内

Fターム(参考) 5B045 BB28 BB42 GG02
5B082 EA01 FA07 HA05 HA08
5K030 LB05 LE03